





Db	1741	ACCCAAAGCAATTCACCGAGCTTGTGAGTTGAGAAGTGCACTGGCTCTGTG	1800	QY	2884	TAGACAGAATGAAATTCACTCAACATTGTTAGATAACGTAGGATTAGAGA	2943
Qy	1804	GCTCAAGCITTATGTTGTTAGAGAACCTTACAGGATCTGATTAACGATTCAG	1863	Db	2881	TAGACAGAATGAAATTCACTCAACATTGTGTTAGATAACGTAGGATTAGAGA	2940
Db	1801	GCTCAAGCITTATGTTGTTAGAGAACCTTACAGGATCTGATTAACGATTCAG	1860	Qy	2944	TTGCCTTAGTGCGCTTTCGCAACTTTCCTTCCTGTTGTTTCTTCGATTTAGG	3003
Qy	1864	CAACTTCAGCAGGCTTAAGGTAAAGCTATGGAAATCTCGCGATGTGTATGATACCG	1923	Db	2941	TTGCCTTAGTGCGCTTTCGCAACTTTCCTTCCTGTTGTTTCTTCGATTTAGG	3000
Db	1861	CAACTTCAGCAGGCTTAAGGTAAAGCTATGGAAATCTCGCGATGTGTATGATACCG	1920	Qy	3004	GTCAGTCAGGCTGAGATTCAGBAGCACTCCAGAGTGGAGCTGAGCTGAGA	3063
Qy	1924	AATTAATGGAGATACTCTGGCTCTAGAAAGGGACTCTGTGACTCTTATAGGC	1983	Db	3001	GTCAGTCAGGCTGAGATTCAGBAGCACTCCAGAGTGGAGCTGAGCTGAGA	3060
Db	1921	AATTAATGGAGATACTCTGGCTCTAGAAAGGGACTCTGTGACTCTTATAGGC	1980	Qy	3064	TATGATACAGTGGCAGAGATACTCTGGCTCTAGAAAGGGACTCTGTGACTCTT	3123
Qy	1984	GTCTGATGATGCGCTATGGTTGGCTTAGAGCTGAGGATCACATATAGAAATCA	2043	Db	3061	TATGATACAGTGGCAGAGATACTCTGGCTCTAGAAAGGGACTCTGTGACTCTT	3120
Db	1981	GTCTGATGATGCGCTATGGTTGGCTTAGAGCTGAGGATCACATATAGAAATCA	2040	Qy	3124	AATGTTACAGAGGTGAGGAATATCACTACANTCAACTAATGTCGAAACTGTG	3183
Qy	2044	GCTTGTGCGAGTTGTTGGAGGATCAACCTGTGACATGATGATCTCCCTGGA	2103	Db	3121	AATGTTACAGAGGTGAGGAATATCACTACANTCAACTAATGTCGAAACTGTG	3180
Db	2041	GCTTGTGCGAGTTGTTGGAGGATCAACCTGTGACATGATGATCTCCCTGGA	2100	Qy	3184	ACATGATTTAGCTGGCCCTGTTGATTCGTTTATAGGTTGGATGGGCA	3243
Qy	2104	CTTGTGAAATGTGTTGGAACACTCTGGCAGGGTTGTTCTAGGTTCAGAGACC	2163	Db	3181	ACATGATTTAGCTGGCCCTGTTGATTCGTTTATAGGTTGGATGGGCA	3240
Db	2101	CTTGTGAAATGTGTTGGAACACTCTGGCAGGGTTGTTCTAGGTTCAGAGACC	2160	Qy	3244	TGCTGAAGGTTGGACTGAGAGCAGCTGAACTGCGCAGCTGGTTGGTTATGAT	3303
Qy	2164	AAGGATAAAAATTAAACTCTGGGACTACTATGATGATCTTGTTTGACTCTTG	2223	Db	3241	TGCTGAAGGTTGGACTGAGAGCAGCTGAACTGCGCAGCTGGTTGGTTATGAT	3300
Db	2161	AAGGATAAAAATTAAACTCTGGGACTACTATGATGATCTTGACTCTTGACTCTTG	2220	Qy	3304	ATACACTGTGAAACTATCTGTGACAGCTGAGCTGAGCTGACAGTCTGAG	3363
Qy	2224	GAAGAGTGAGGTTAGTCAGGGTTCTCTTGTGTCGTCGACTATGGAGATT	2283	Db	3301	ATACACTGTGAAACTATCTGTGACAGCTGAGCTGACAGTCTGAG	3360
Db	2221	GAAGAGTGAGGTTAGTCAGGGTTCTCTTGTGTCGTCGACTATGGAGATT	2280	Qy	3364	TGCTGAAGGCAACTCTGGGACTCTGCTGTCGTTATCTGAGCTGAGCTGAG	3423
Qy	2284	GGAGCCGAGCATGTGAAAGCTAGTGTGATGAGGAGCTGGAGCTGAG	2343	Db	3361	TGGTGAAGGCAACTCTGGGAGCTGCTGTCGTTATCTGAGCTGAG	3420
Db	2281	GGAGCCGAGCATGTGAAAGCTAGTGTGATGAGGAGCTGGAGCTGAG	2340	Qy	3424	ATGCTACTGTGAGAACTACACACAGATACGAGCTGAGCTTCTGGCCAGGT	3483
Qy	2344	TATACAGAGTAGAAACTCTGGCCCTGATGTTGACGCTTGTGAGCTGAG	2403	Db	3421	ATGCTACTGTGAGAACTACACAGATACGAGCTGAGCTTCTGGCCAGGT	3480
Db	2341	TATACAGAGTAGAAACTCTGGCCCTGATGTTGACGCTTGTGAGCTGAG	2400	Qy	3484	GGAAATACCTGTGAGGCTCTGTCGTCGATCATAPATACTCATATGAG	3543
Qy	2404	CCTGTTGTTACATGTGAGGCTGATGTTGACGCTTGTGAGCTGAGCTA	2463	Db	3481	GGAAATACCTGTGAGGCTCTGTCGTCGATCATAPATACTCATATGAG	3540
Db	2401	CCTGTTGTTACATGTGAGGCTGATGTTGACGCTTGTGAGCTGAGCTA	2460	Qy	3544	CTTCGGAGTTCTCTGTTCTGTAATCTCTCTGAGTTAGTGTATAATGACA	3603
Qy	2464	AGGCCCTGTGAAACTTGTGAACTATGATGTTGACGAGGCTCTGAGGT	2523	Db	3541	CTTCGGAGTTCTCTGTTCTGTAATCTCTGAGTTAGTGTATAATGACA	3600
Db	2461	AGGCCCTGTGAAACTTGTGAACTATGATGTTGACGAGGCTCTGAGGT	2520	Qy	3604	CAAAATAATGCTTGGCACACCTTCTGAGTAACTATACATAAGGCT	3663
Qy	2524	AGGGTTGTTGAAACTCTGGCTGTTGTGATGTTGAGGAGCTGAGCTG	2583	Db	3601	CAAAATAATGCTTGGCACACCTTCTGAGTAACTATACATAAGGCT	3660
Db	2521	AGGGTTGTTGAAACTCTGGCTGTTGTGATGTTGAGGAGCTGAGCTG	2580	Qy	3664	ACAA 3667	3661
Qy	2584	AGCATCTCTGAGCTCTGGCTGGAATGCTGTCAGGAGAGTATT	2643	Db	3661	ACAA 3664	
Db	2581	AGCATCTCTGAGCTCTGGCTGGAATGCTGTCAGGAGAGTATT	2640				
Qy	2644	CTTAAAGGAGCTCATCTTCTGAGGAGTGGTTCTCTATGGATGTGTC	2703				
Db	2641	CTTAAAGGAGCTCATCTTCTGAGGAGTGGTTCTCTATGGATGTGTC	2700				
Qy	2704	GCTTACCATAGTTGATTAATGATGCAATTCTATATCTGATGTC	2763				
Db	2701	GCTTACCATAGTTGATTAATGATGCAATTCTATATCTGATGTC	2760				
Qy	2764	TGTTTGTGAGGTTAGACATGTCACCTTATACATGTC	2823				
Db	2761	TGTTTGTGAGGTTAGACATGTCACCTTATACATGTC	2820				
Qy	2824	ATTACAGTGTGAGTAATTCACTAATTGCTGTGAGTTTGATGAACTG	2883				
Db	2821	ATTACAGTGTGAGTAATTCACTAATTGCTGTGAGTTTGATGAACTG	2880				
		RESULT 2					
		AY221467					
		LOTUS					
		AY221467					
		DEFINITION					
		Arabidopsis thaliana truncated division protein ('arc6') gene, arc6-1					
		allele; complete cds; nuclear gene for chloroplast product.					
		ACCESSION					
		AY221467					
		VERSION					
		AY221467.1					
		KEYWORDS					
		'					
		SOURCE					
		Arabidopsis thaliana (thale cress)					
		ORGANISM					
		Arabidopsis thaliana					
		Arabidopsidae					
		Rosidae					
		Viridiplantae					
		Spermatophyta					
		Magnoliophyta					
		eudicotyledons					
		core eudicots					
		rosids; euRosids II; Brassicales; Brassicaceae; Arabidopsis					
		REFERENCE					
		Wittha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Erp,H. and					
		AUTHORS					



Db	1501	GAAGAAGCTTAATGTTTAAACGGGTGCGGAAATTGTCGCTGTTGAGGAGGTGG 1560
Qy	1564	AGCATCACTCTTGTGAGGGTTTCAACCGTGAGAAGTTTATGATGGCGTTAACG 1623
Db	1561	AGCATCACTCTTGTGAGGGTTTCAACCGTGAGAAGTTTATGATGGCGTTAACG 1620
Qy	1624	ATGCACTGCTGAGCAGGTTACAGTTAGTACCTTTTAATTCTTAGCATGA 1683
Db	1621	ATGCACTGCTGAGCAGGTTACAGTTAGTACCTTTTAATTCTTAGCATGA 1680
Qy	1684	TATRACTTGTGTTCTGATTATGATGTTGTTGTTGTTGAGTTGAACTTTTGTAGT 1743
Db	1681	TATRACTTGTGTTCTGATTATGATGTTGTTGTTGTTGAGTTGAACTTTTGTAGT 1740
Qy	1744	ACCCCAAGCAATTTCAGCAGAGTCATTGAGTTTAACTGACTGCTGCTG 1803
Db	1801	GCTCAAGCTTATGTTGTTAGAGACCACCTTACAGGATCTGATAGCATTCAG 1860
Db	1741	ACCCCAAGCAATTTCAGCAGAGTCATTGAGTTTAACTGACTGCTGCTG 1800
Qy	1864	CACTTCUGCAGCTTAAGGTTAGAGCTTACAGGATCTGAGTTTAACTGACTGCTG 1923
Db	1861	CAACTTCUGCAGCTTAAGGTTAGAGCTTACAGGATCTGAGTTTAACTGACTGCTG 1920
Qy	1924	AATTAATTGGAGATAGACTTCGCTCTAGAAAGGGACTCTGCACTGCTTATAGGCAA 1983
Db	1921	AATTAATTGGAGATAGACTTCGCTCTAGAAAGGGACTCTGCACTGCTTATAGGCAA 1980
Qy	2044	GCTTATGGGACTTTGGAGATTCAACTGATGACATGATCTCCITGA 2103
Db	2041	GCTTATGGGACTTTGGAGATTCAACTGATGACATGATCTCCITGA 2100
Qy	2104	CTTGCAATTGTTGGAAACCTGGTGGAGGTTGCTTCTAGGTTCAGGACCC 2163
Db	2161	AAAGATAAAATTAAACTCGGGACTACTATGATGCTCTAGGTTCAGGACCC 2220
Db	2101	CTTGCAATTGTTGGAAACCTGGTGGAGGTTGCTTCTAGGTTCAGGACCC 2160
Qy	2224	GAAGAGCTGGGAGTAGTTGGGTTCTCTTGTGCTGCTGCTGCTGCTGCTG 2283
Db	2221	GAAGAGCTGGGAGTAGTTGGGTTCTCTTGTGCTGCTGCTGCTGCTGCTG 2280
Qy	2284	GGGCCGACCATGAGAACTCTGCTATGCAAGCACTCGAGAAGTTCTTCGCGC 2343
Qy	2281	GGGCCGACCATGAGAACTCTGCTATGCAAGCACTCGAGAAGTTCTTCGCGC 2340
Qy	2344	TATACAGATAGAACTCTGCTGAGGTTGCTGCTGCTGCTGCTGCTGCTG 2403
Db	2341	TATACAGATAGAACTCTGCTGAGGTTGCTGCTGCTGCTGCTGCTGCTG 2400
Qy	2404	CCTCTGTGAACTACATGAGCTGCTGAGGTTGCTGCTGCTGCTGCTGCTG 2463
Db	2401	CTTCTGTGAACTACATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Qy	2464	AGGCCTCTGAAACTACTGTGAACTATGATGATGAACTGAGCTGGGCTCTGAGCT 2523
Db	2461	AGGCCTCTGAAACTACTGTGAACTATGATGATGAACTGAGCTGGGCTCTGAGCT 2520
Qy	2524	AGCTCTGAAACTACTGTGAACTATGATGATGAACTGAGCTGGGCTCTGAGCT 2583
Db	2521	AGCTCTGAAACTACTGTGAACTATGATGAACTGAGCTGGGCTCTGAGCT 2580
Qy	2584	AAATCTCTGCTGGGCTGGCAATTGCTACTGTGAGCTGGCAAGAGTATTG 2643

Db	2581	AAGATCTCTAGCTCTGGGTTGGCAATTGACTGATTCACGTCAGCCAGAGTATT 2640
Qy	2644	CTTAAAGGAGCTCATCTTTCATGCGAGGATATGGTTCTCTATGGAATCTGTC 2703
Db	2641	CTTAAAGGAGCTCATCTTTCATGCGAGGATATGGTTCTCTATGGAATCTGTC 2700
Qy	2704	GCTACCATAGGTAGTAACTGCAATTCTCATATTCCTGATGCTCAAATGTC 2763
Db	2701	GCTACCATAGGTAGTAACTGCAATTCTCATATTCCTGATGCTCAAATGTC 2760
Qy	2764	TGTTTGTGAGCTAGAGCATAGTCCACTTAATACATGTCGAAAGTGTACAG 2823
Db	2761	TGTTTGTGAGCTAGAGCATAGTCCACTTAATACATGTCGAAAGTGTACAG 2820
Qy	2824	ATPACACAGTGTGCTGAACTTTCACTPATATGCTGCTGTTGATGCAACTG 2883
Db	2821	ATPACACAGTGTGCTGAACTTTCACTPATATGCTGCTGTTGATGCAACTG 2880
Qy	2944	TGCTCTAGTGTGCTGCTGCTGCACTTTCTTCTGTTGATTTCTTCGTTGATTAG 3003
Db	2941	TGCTCTAGTGTGCTGCTGCTGCACTTTCTTCTGTTGATTTCTTCGTTGATTAG 3000
Qy	3004	GTAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3063
Db	3001	GTAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3060
Qy	3064	TATGATTCAGTGGCAAGATACTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTG 3123
Db	3061	TATGATTCAGTGGCAAGATACTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTG 3120
Qy	3124	AATGTTACACAGCTGAGGAAATACTACATCAATCAATCAATCAATCAATGTTGAAACTGTG 3183
Db	3121	AATGTTACACAGCTGAGGAAATACTACATCAATCAATCAATCAATGTTGAAACTGTG 3180
Qy	3184	ACATGATTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3243
Db	3181	ACATGATTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
Qy	3244	TGCTGAAGTTGACTGACAGGAGCTGAGCTGAACTGCGCAGCTGGTTGTTGATG 3303
Db	3241	TGCTGAAGTTGACTGACAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 3300
Qy	3304	ATACACTGTGAAACTATCTGTTGACAGTGTGAGCTGCTGAGATGGAAACCGTGC 3363
Db	3301	ATACACTGTGAAACTATCTGTTGACAGTGTGAGCTGCTGAGATGGAAACCGTGC 3360
Qy	3364	TGCTGAAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3423
Db	3361	TGCTGAAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420
Qy	3424	ATGCTACTGATGCTGAGCTACACAGAGTACAGGTTTCTGCTGCTGCTGCTGCTG 3483
Db	3421	ATGCTACTGATGCTGAGCTACACAGAGTACAGGTTTCTGCTGCTGCTGCTGCTG 3480
Qy	3484	GGAAATACATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3543
Db	3481	GGAAATACATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3540
Qy	3544	CTTCTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3603
Db	3541	CTTCTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600
Qy	3604	CAAAAATTAACCTCTGGCACACCTTCTCTGATCTAACTACATACATAGGCT 3663
Db	3601	CAAAAATTAACCTCTGGCACACCTTCTCTGATCTAACTACATACATAGGCT 3660
Qy	3664	ACAA 3667
Db	3661	ACAA 3664

RESULT 3  
 AB016888 AB016888 85791 bp DNA linear PIN 14-FEB-2004  
 LOCUS Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MDH9.  
 DEFINITION  
 ACCESSION AB016888 BA000015  
 VERSION AB016888.1 GI:3449329  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE ORGANISM  
 REFERENCE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 AUTHORS Asamizu, E., Sato, S., Kaneko, T., Nakamura, Y., Kotani, H., Miyajima, N.  
 TITLE Structural analysis of Arabidopsis thaliana chromosome 5. VIII.  
 JOURNAL Sequence features of the regions of 1,081,958 bp covered by  
 severenteen physically assigned P1 and TAC clones  
 MEDLINE DNA Res. 5 (6), 379-391 (1998)  
 PUBLISHER 9915233  
 REFERENCE 2 (bases 1 to 85791)  
 AUTHORS Nakamura, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research  
 INSTITUTE, Department of Plant Gene Research, 1532-3, Yana,  
 Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,  
 Tel:01-438-52-3935, Fax:01-438-52-3934)  
 Address for correspondence: kaos@kazusa.or.jp  
 For the latest information on annotation of this clone, please see  
 http://www.kazusa.or.jp/kaos/cgi-bin/agd\_graph.cgi?c=MDH9  
 Genes with similarity to proteins in the databases are described in  
 'product' or 'note' qualifiers. Genes that have no significant  
 protein similarity are described as 'unknown protein'.  
 The software programs used to predict genes include: Grail  
 (Informatics Group, Oak Ridge National Laboratory,  
 http://compbio.ornl.gov/Grail-1.3/), CCR-081.mitt.edu/GENSCAN.html,  
 GENSCAN (Chris Burge, MTR, http://CCR-081.mitt.edu/GENSCAN.html),  
 NetGene2 (S.M. Hebsgaard et al., 'Technological University of  
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/), and  
 SplicePredictor (Volker Brendel, Stanford University,  
 http://grail.cse.zool.usask.ca/~edu/cgi-bin/sp.cgi).  
 Genes encoding tRNAs are predicted by tRNAscan-SE  
 (Sean Eddy, Washington University School of Medicine, St. Louis,  
 http://genome.wustl.edu/~eddyl/tRNAscan-SE/).  
 This sequence may not be the entire insert of this clone. It may be  
 shorter because we remove overlaps between neighboring submissions.  
 The 5' clone is K5114 and the 3' clone is K1681.  
 Location/Qualifiers  
 1. .85791  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3702"  
 /chromosome="5"  
 /clone="MDH9"  
 /clone\_lib="Mitsui PI"  
 /ecotype="Columbia"  
 /complement("59...239")  
 /note="CDS 18 reported in Acc# AB023032  
 gene\_id:K5114..12  
 unknown protein"  
 /number=1  
 /evidence="not\_experimental  
 complement (Join(2056..2079,2278..2393,2991..3633))  
 /note="contains similarity to unknown protein  
 gene\_id:MDH9..1  
 pir|T04864"  
 /pseudo  
 /codon\_start=1  
 /evidence="not\_experimental  
 join(5739..6138,6462..6595)

FEATURES  
 Source

CDS  
 /notes="unnamed protein product; gene\_id:MDH9..2  
 /protein\_start=1  
 /evidence="not\_experimental  
 /protein\_id="BAB10474..1"  
 /db\_xref="GI:9759472"  
 /translation="MISEKPLGEESIRODLEVLTVSRLLVKSYQSKLKKKHTEV  
 DEEEYARGAVCNCLISVGCRVADGFDSSDNRSRWAASEDESGKLMWTCBTRL  
 DCFSYGVREPKWKKNRYLAQSDGYRKHVDPPDILCMMLPLPSLNLAMBLCK  
 KQSQMANTORFLQRERGSQTIPWFLFAALKDGCGSSGGIHHGIVDQSQRHRFTDL  
 KGKFMYSVTSHEBITYGRSMDRNSPFSKSHRGVLSVDFPRLSYRRONSNDAGTK  
 SHRLRKQDRNURKNSKRVFVLAIGGTGFDEFJUDGEYTSATNTNSMQJRMGP  
 GVVSQGICINGFVAYSENQDGLSGIDIERGFWTTSPSPRPPRHEFYQKLVSCHRL  
 FMSYSWCDDGGDQVYGRKNAVKWLWDLVYLTWTEVSPHPDAPMDPQFVDPYQKSCWHL  
 LMGLEMFKFGQVISFFTCVDLTEASWHRHSRQSQKLNSCTKNTLHL"  
 complement(15087..16778)  
 /note="unnamed protein product; gb|AAFO780..1  
 gene\_id:MDH9..5  
 similar to unknown protein"  
 /codon\_start=1  
 /evidence="not\_experimental  
 /protein\_id="BAB10477..1"  
 /db\_xref="GI:9759472"  
 /translation="MISEKPLGEESIRODLEVLTVSRLLVKSYQSKLKKKHTEV  
 DEEEYARGAVCNCLISVGCRVADGFDSSDNRSRWAASEDESGKLMWTCBTRL  
 DCFSYGVREPKWKKNRYLAQSDGYRKHVDPPDILCMMLPLPSLNLAMBLCK  
 KQSQMANTORFLQRERGSQTIPWFLFAALKDGCGSSGGIHHGIVDQSQRHRFTDL  
 KGKFMYSVTSHEBITYGRSMDRNSPFSKSHRGVLSVDFPRLSYRRONSNDAGTK  
 SHRLRKQDRNURKNSKRVFVLAIGGTGFDEFJUDGEYTSATNTNSMQJRMGP  
 GVVSQGICINGFVAYSENQDGLSGIDIERGFWTTSPSPRPPRHEFYQKLVSCHRL  
 FMSYSWCDDGGDQVYGRKNAVKWLWDLVYLTWTEVSPHPDAPMDPQFVDPYQKSCWHL  
 complement(Join(18111..18193,18281..18535,18773..19001,  
 19088..19195,19307..19447,1952..1952,20129..20245))  
 /note="unnamed protein product; gene\_id:MDH9..6  
 similar to unknown protein  
 sp|P42251"  
 /codon\_start=1  
 /evidence="not\_experimental

CDS  
 CDS

CDS

			/protein_id="BAB10479_1"
		/ab_xref="GI:9759473"	
		/translation="MNOISFNGLSSLILUTVIAABEPITRITAFGSCANQSAOPTINDA	
		INKFDPLITWGDNYGDPDRLPRTIGKERTFGPRNSPRPVSPSEREMLYAK	
		ANPGYRIOQRKAVIGHDDHGYGLDAGKEDRKVNOKNQMLDPEPNEPLSPRQA	
		GTVASTTYGPPRKKVUDPQKVKVLLTIVDGVPIGSVHFGIT	
		RYDCSVCYPLVNTVSILVQSVEKVVRPLSI VRULFWYTPSTMVINDCKPKCT	
		YCOONIGRAISDNANATVTKLIRIDYNGTIVNLSLQPGSSNSLKDTTGGK	
		SORYCULIURGTRTRVHLVLTIVTAVLAMALGJLIGAVLAATRCVCKVD",	
CDS		complement(21540 .. 22097)	
		/note="unnamed protein product; contains similarity to	
		calmodulin	
		gene_id:MDH9_7."	
		/codon_start=1	
		/evidence=not experimental	
		/protein_id="BAB10479_1"	
		/ab_xref="GI:9759474"	
		/translation="MNLAKNQKSLSLRLKKVSSKOSSESSRNLEDERTSSNSSSS	
		LNVNERARTVFDMDANSKGKLSGSEQCSVSLGGLALSREVEVKTSVDGDK	
		KUMIREPDQNDGUSPDEFULMR"	
CDS		Join(24090 .. 24289, 2410) .. 24614, 24701 .. 24945, 25115 .. 25313,	
		25632 .. 25679, 25821 .. 25938, 26039 .. 26227, 26314 .. 26613,	
		26730 .. 26931, 27169 .. 27353, 27487 .. 2771, 27939 .. 28122,	
		28213 .. 28380, 28646 .. 28735, 28889 .. 29026, 29204 .. 29431,	
		29519 .. 29671, 29784 .. 29915, 30041 .. 30554 .. 30674,	
		30763 .. 30943, 31932 .. 31033, 31112 .. 31262, 31338 .. 31429)	
		/note="gene_id:MDH9_8"	
		/codon_start=1	
		/evidence=not experimental	
		/product="pitri-lysin"	
		/ab_xref="GI:9759475"	
		/translation="MASSSSSTFTGKFSPITALPFGSDRSRVRKYLDSRNKVRKFPSC	
		CKRNQGTRRLPSA9DRTFPLSLSLRSKRSQIYWA7GDPFHAGTAWD	
		GIVAEQQDLDLUPPBDGABEAGLFCPLSPKPKRQKQGKLRILILPKVPRF	
		BHMFETHGSDTDEBEGDQIAMIYTFALVAFPLKPSKRRKLLTGARSNAYTDFHTPHFI	
		HSPTKTDSDLPSPVLDALNEIAPHKPEKSRVKEERAISELQMMVTEYDQ	
		OLOLHQJSENKGRKPFGLIGERBKQKMDVKRKFWRWYDIDNIPR	
		IWHNTRAVFGKQDNLBTESTSPSPRASRQVAMANFLPFLKPLAGLGFNSNCINTDQS	
		KMTKRHRATRPPVHNWLSPGTLKPPDPIKHQLQIAINMECKIPIKSVKPTG	
		DLRNVMKRTFLSALPHRINTKRNSSPMLVPSVGLHDSRREGCVTHTLIPKVRP	
		WONAIVKAVQEVPLRLLKEFGVTRGELRYYDPMVLSKSHLAMIDNNSVLUDEMES	
		DALSHTMDOPOGHETFLWAGVTFYEVNTGAKLPTESDCEPRTAIPAAV	
		PTKVHFDVGSESDNPDTELESVSGSLIATCAVNYKSTTRESRAGVPLKELIQ	
		NPCFVUIGPSGSKLKDENTCITOLQKQVYKQVYKQVYKQVYKQVYKQVYKQ	
		TSDSKQAWVVGRTLSSEGGRVGDPSQEVLRFVNHLINGLSESTERIAMEPFLR	
		DNGMQAQFOLQHMLVRSVWLEDAFRAROLXLYSYRSPISLERATVAKLIMANG	
		DERFVPTPKSLOSUNLVEKSDAUMHVGUVGOMENVSYDGESEBIRCLDYLGVK	
		ASHDSKPPGSPBPLRQTAGLQFOPVFLDTERBCVAGFARRWMTGDDLI	
		FQSVSLUPVANDGLLSSBQQLGKRRBLQKGRKMLPFFTMGLIAETNSRIFTT	
		VRDSLQIUTYDVSPELFLDPLWVYVTSPTPSKVYKAVDCKYVGHNSQDAP	
		RELDRIKRKLMLRHELKSWAYWNLAHQASSPRKELSCIKELSVYEASIED	
		ITYLAVNQRLVDEDLSLSCIGIAGAQGBEITWLSEBEBPEDVPSGVVPVERGSSMTR	
		PTT"	
CDS		complement(Join(33251 .. 33326, 33512 .. 33663, 33744 .. 33797,	
		Query Match	
		Best Local Similarity 99.8%; Score 3659; DB 8; Length 85791;	
		Matches 3662; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
Qy	1	TGTCTGCTATTAGGGAGATACTATTATAGCACTTGTCTGTTACATAGATTTCG	60
Db	63597	TGTCTGCTATTAGGGAGATACTATTATAGCACTTGTCTGTTACATAGATTTCG	63656
CDS		complement(Join(33251 .. 33326, 33512 .. 33663, 33744 .. 33797,	
		Query Match	
		Best Local Similarity 99.8%; Score 3659; DB 8; Length 85791;	
		Matches 3662; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
Qy	1	TGTCTGCTATTAGGGAGATACTATTATAGCACTTGTCTGTTACATAGATTTCG	60
Db	61	TGTCTGCTATTAGGGAGATACTATTATAGCACTTGTCTGTTACATAGATTTCG	120
Db	63657	TGTCTGCTATTAGGGAGATACTATTATAGCACTTGTCTGTTACATAGATTTCG	63716
Qy	121	AATTACACATTTCAGTGGATGTTAGAAGAAGAGCGGAATGTGGGTGTTGGGG	180
Db	6397	AATTACACATTTCAGTGGATGTTAGAAGAAGAGCGGAATGTGGGTGTTGGGG	63776
Qy	181	TTAACATTAAAGTAGTCAGTCAAATAGTTCTGCTCTATATGTTCA	240
Db	63777	TTAACATTAAAGTAGTCAGTCAAATAGTTCTGCTCTATATGTTCA	63836
CDS		241 ATGAGGGCAACTCAAACGGTCTAGGGAAATAGATTATGGACGTTAACATG	300
		63837 AATGAGGGCAACTCAAACGGTCTAGGGAAATAGATTATGGACGTTAACATG	63896
		301 AAATCCGGGGCAACTGTTAGCATAGTGTGTTATTTTCAGTAGCTGGT	360
Db	63897	AAATCCGGGGCAACTGTTAGCATAGTGTGTTATTTTCAGTAGCTGGT	63956
Qy	361	GTWTGTTGTTGTTACTTAACTTAACTCAAATTCATAACCCATAGGACAA	420
Db	63957	GTWTGTTGTTGTTACTTAACTCAAATTCATAACCCATAGGACAA	64016
Qy	421	CAGTCCTCTCAATPTGAAACAGACAGTTTGAGCTAAAGACACTCC	480
Db	64017	CAGTCCTCTCAATPTGAAACAGACAGTTTGAGCTAAAGACACTCC	64076
Qy	481	ATGAGAGCTCTGACTCACCTGCCTGCTCTCCCAATTCCATATGCGT	540
Db	64077	ATGAGAGCTCTGACTCACCTGCCTGCTCTCCCAATTCCATATGCGT	64136
Qy	541	CCGGCGGAGACAAGCTCCGCTAGGCCACAAACCCCTACACTATGCT	600
Db	64137	CCGGCGGAGACAAGCTCCGCTAGGCCACAAACCCCTACACTATGCT	64196
Qy	601	AAATGGGGCACCTCTCTCGACACTCAATTCACTCCGATTCCTCCCTCC	660
Db	64197	AAATGGGGCACCTCTCTCGACACTCAATTCACTCCGATTCCTCCCTCC	64256
Qy	661	TTGCCACGCCACACCACCCGCACTCTGTCCTCTGCCACCATCTATG	720
Db	64257	TTGCCACGCCACACCACCCGCACTCTGTCCTCTGCCACCATCTATG	64316
Qy	721	GAACGCCAGTCCTCCATCCCATGTTCTACCGGTTAGGAGCTAACACATT	780
Db	64317	GAACGCCAGTCCTCCATCCCATGTTCTACCGGTTAGGAGCTAACACATT	64376
Qy	781	TTACCGATGGATCTAGAGACATTAGTCAGTGGTTGAGCTTCGAACCGCAATTG	840
Db	64377	TTACCGATGGATCTAGAGACATTAGTCAGTGGTTGAGCTTCGAACCGCAATTG	64436
Qy	841	TTACGGAGGACCTTAAATCACCGGAGACAGATTCTTCAACTGCTGCTG	900
Db	64437	TTACGGAGGACCTTAAATCACCGGAGACAGATTCTTCAACTGCTGCTG	64496
Qy	901	TCTAATCCCTGGCTCTAGAGAGTCACTGAGTCGTTCTGATGAGAGCTAA	960
Db	64497	TCTAATCCCTGGCTCTAGAGAGTCACTGAGTCGTTCTGATGAGAGCTAA	64556
Qy	961	CTACATCACTGAGCTCTGGATGAGAGTCACTGAGTCGTTCTGATGAGAGCTAA	1020
Db	64557	CTACATCACTGAGCTCTGGATGAGAGTCACTGAGTCGTTCTGATGAGAGCTAA	64616
Qy	1021	GTTTAATTCTAGTAATGTTAGAAGAGGAACATTATCTAGTGATGAGGTC	1080
Db	64617	GTTTAATTCTAGTAATGTTAGAAGAGGAACATTATCTAGTGATGAGGTC	64676
Qy	1081	TCTACGTTGTTGCTCTAGAGAGTCACTGAGTCGTTCTGATGAGAGCTAA	1140
Db	64677	TCTACGTTGTTGCTCTAGAGAGTCACTGAGTCGTTCTGATGAGAGCTAA	64736
Qy	1141	GCTTAAGGAGGCTGCTACTGAGATGAGTCGTTCTGATGAGAGCTAA	1200
Db	64737	GCTTAAGGAGGCTGCTACTGAGATGAGTCGTTCTGATGAGAGCTAA	64796
Qy	1201	GTTCCTCGAGTCTCGAGGAGACATGGATGAGTCGTTCTGATGAGAGCTAA	1260
Db	64797	GTTCCTCGAGTCTCGAGGAGACATGGATGAGTCGTTCTGATGAGAGCTAA	64856



Db 67017 ACGATGCTACTGAGTCAGACAGATACAGATCAGTGTTCAGTG 67076

Qy 3481 GGTTGAAATCTGAAAGCTCTGCAATATACATGATGAGCATGCT 3540

Db 67077 GGTGAAATCTGAAAGCTCTGCAATATACATGATGAGCATGCT 67136

Qy 3541 GAGCTTGAGATTCCTCTGTCTGTAATTCTCTCTAAGTAGTTATAATGA 3600

Db 67137 GAGCTTGAGATTCCTCTGTCTGTAATTCTCTCTAAGTAGTTATAATGA 67196

Qy 3601 ACCAAAAATTAACTACGTTGGCACCCCTTCCTGATTAACATACATAAGG 3660

Db 67197 ACACAAAAATTAACTACGTTGGCACCCCTTCCTGATTAACATACATAAGG 67256

Qy 3651 GCTCAA 3667

Db 67257 GCTCAA 67263

RESULT 4  
AY091075

LOCUS AY091075 mRNA linear PLN 18-SEP-2002

DEFINITION Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete

ACCESSION AY091075

VERSION AY091075.1 GI:20259550

KEYWORDS F1I; CDNA; Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana (thale cress)

Bukaryota; Viriplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis;

(bases 1 to 2637)

REFERENCE Yamaoka, K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bower,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shimizu,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R., and Theologis,A.

TITLE Arabidopsis Full Length cDNA Clone

JOURNAL Unpublished

AUTHORS Yamaoka, K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,

Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,

Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bower,L.,

Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,

Sakurai,T., Satou,M., Seki,M., Shimizu,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R., and Theologis,A.

JOURNAL Submitted (21-MAR-2002) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y., and Shinozaki,K.

The Salk, Stanford, PGBC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bower,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shimizu,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGBC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGBC)

contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis Genome submitted to GenBank.

#### FEATURES

##### Source

1. 2037

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="Taxon:3702"

/chromosome="5"

/clone="RAT09-76-G11 (R19395)"

/ecotype="Columbia" (FLC-1) as a BamHI/XbaI insert."

##### gene

##### 5' UTR

CDS

/gene="At5g42480"

114..2519

/gene="At5g42480"

/codon\_start=1

/evidence=experimental

/product="unknown protein"

/protein\_id="AAM13895.1"

/db\_xref="GI:20259551"

/translation="MEALSHVIGLISPFOLCPRLPATKURRSHTSTICSSASKNAD

RULSDERFTSOSSSRSPATATTATVSPSSPSIDRERHVPVPIDFVQLGQTHFJ

DTGTRAFEARVSKPPGFSDDALISRRQIQAACETNSNRSRRENEGLDEBT

VITDVEDPKVOCALCQDGEGETTEVLRKGKLLERKLESFKDWVMLAFLDVS

RKDAWALDPDFDTGYSVEREAKLKLQDPLAQRQDTEBTLTIPRVLBEG

LPLGDYDAKRNGLSGVRNTLWSVGGSASLVLGULTRKPNTEFLRMTAEQD

FVATPSNIPASSEPEVVALEVALVAQFCKKHHQDADKOFQLOQAKUMAMEIPM

FVATPSNIPASSEPEVVALEVALVAQFCKKHHQDADKOFQLOQAKUMAMEIPM

LYDTTRNWEIDPGLERLICALLQGKDVCRWNLGLOSDSRNPALVEFLJENSRD

DNDLGCLKLLTETWLAGWVPERFRDKKPKLQYDDPMVLSLERYVNGQSL

AKAAAMARKIGAKRVSASQMLQKVPSYRDKRRAETIQOLQYDPMVLSLERYVNGQSL

GRGVFTIAEVTPRSENFTENDYAIRGVSESSVDETMVSAWMIKEASRFLKLAQV

AGLISLUFSDOKFLKSSRKFDRMWSMSBSDVATGVSRADESDLPROMARTANI

VSKWQKTKSLRGPDIRMLEFEDUGRMKLUWDRRAETIQOLQYDPMVLSLERYVNGQSL

VTVMSADQTALVEATLERESACSLDLVPHENNADVRYTRYEVFWSKSGWKRTEGSV

LAS"

2537..2637

#### ORIGIN

Query Match 52.7%; Score 1931; DB 8; Length 2637;  
Best Local Similarity 80.4%; Pred. No. 0; Mismatches 0; Indels 5; Gaps 5;

Matches 2616; Conservative

1 GATTAACTTACTCTAACATTCAAACTCCATTAACTCCAGGACCAACAGTC

368 GATTAACTTACTCTAACATTCAAACTCCATTAACTCCAGGACCAACAGTC

1 GATTAACTTACTCTAACATTCAAACTCCATTAACTCCAGGACCAACAGTC

428 TTCAATATCTAAACAGAACAAAGTTTGTAGCTAAAGACACTCCATGGAG

61 TCAATATCTAAACAGAACAAAGTTTGTAGCTAAAGACACTCCATGGAG

489 CTCTGAGTCAGTCGGCATGTCCTCCATCTCCGATTCACCCGATTCCTCC

121 CTCTGAGTCAGTCGGCATGTCCTCCATCTCCGATTCACCCGATTCCTCC

548 CGACAACTCCGAGTACACACTCTAACATCTCTCCGAGAAATCG

181 CGACAAGTCGAGTCGGCATACACTCTACATCTCTCCGAGAAATCG

608 CGACCGGTTCTCCGACTCTCTCTGCACCACTCTATGATGTCCTCCGAGCG

241 CGACCGGTTCTCCGACTCTCTCTGCACCACTCTATGATGTCCTCCGAGCG

668 CGGCACCCACGCCACTCTCTGCACCACTCTATGATGTCCTCCGAGCG

301 CGGCACCCACGCCACTCTCTGCACCACTCTATGATGTCCTCCGAGCG

728 ACGTCCTCATCCCCATGATTCTACCAAGGTTAGGAGCTAACACACCTCTAACCG

787

Db	361	ACGTCCCCATCCGATTGATTCTTACCGGTATAGGACTCAACACITTTAACG	420	Db	1189	AACCTTTATGGTAAGAACCCATTACAGGAATGCTGATACAACTTCAGAAC	1248
Qy	788	ATGGRATCAGAAGAGCATTCGAGACTTAGGGTTGAAACCGCCCAATTGGTTAACGG	847	Qy	1868	TTCAGCAGCTAAGTAATGGATAGGAGATTCTGGGATGTGATGATACAGGATA	1927
Db	421	ATGGRATCAGAAGAGCATTCGAGACTTAGGGTTGAAACCGCCCAATTGGTTAACGG	480	Db	1249	TTCAGCAGCTAAGTAATGGATAGGAGATTCTGGGATGTGATGATACAGGATA	1308
Qy	848	ACGAGCTTAAATGCCGAGAGATCTTCAGCTCTTGAAACCGCCCAATTGGTTAACGG	907	Qy	1928	ATTCGGAGATAGACTTCGGCTGAAAGGGGACTCTGCACTGCTTATGGCAAGTG	1987
Db	481	ACGAGCTTAAATGCCGAGAGATCTTCAGCTCTTGAAACCGCCCAATTGGTTAACGG	540	Db	1309	ATTCGGAGATAGACTTCGGCTGAAAGGGGACTCTGCACTGCTTATGGCAAGTG	1368
Qy	908	CTGGCTCTAGAAGAGAGTCAAGAGGCTTCTTGATGATGAGAAGGCTACGTCA	967	Qy	1988	ATGATGCGCTAATGGGGCTTACAGCAGGATTCAAAATAGAATTCAGGATCCAGCA	2047
Db	541	CTGGCTCTAGAAGAGAGTCAAGAGGCTTCTTGATGATGAGAAGGCTACGTCA	600	Db	1369	ATGATGCGCTAATGGGGCTTACAGCAGGATTCAAAATAGAATTCAGGATCCAGCA	1428
Qy	968	CTTGGTCTCTGGATAGGTTAATTCGATTGGATAATAAAGTTCTGTTAA	1027	Qy	2048	TTGGGAGCTTGGAGAATTCAGTGTGATGATGATCTCCOTGGACT	2107
Db	601	CTTGGTCTCTGGATAGGTTAATTCGATTGGATAATAAAGTTCTGTTAA	617	Db	1429	TTGGGAGCTTGGAGAATTCAGTGTGATGATGATGATGATGATGATGATGATGAT	1488
Qy	1028	TTTCATGAATTGATAAAGGAAGGAACTTTACTAAGAAGGTTCCMGGGCTCTCT	1087	Qy	2108	GCAATTGTTGGAAACCTGGTGGAGGGTGTCTTCCTAGGTCTAGAGCACCAA	2167
Db	618	-----	638	Db	1489	GCAATTGTTGGAAACCTGGTGGAGGGTGTCTTCCTAGGTCTAGAGCACCAA	1548
Qy	1088	GTATTCGAGAAGGTTGGAGAGATGAGATAGTCTCGGGTTGGAGCTCTGCTT	1147	Qy	2168	ATRAAAATTAAACTCGGGGACTACTATGATGATCTCTATGTTTGAGTTAC	2227
Db	639	GTATTCGAGAAGGTTGGAGAGATGAGATAGTCTCGGGTTGGAGCTCTGCTT	698	Db	1549	ATRAAAATTAAACTCGGGGACTACTATGATGATCTCTATGTTTGAGTTAC	1608
Qy	1148	GAGAGGTTGCCTAATGCTGTTAAGCAAGAGTGCTGTTAGTTAGGCTTGCTTC	1207	Qy	2228	GACTGGGGAGTTGGAGCTTGTGACTGCTTGGGTTAGTGTAGATCTGAA	2287
Db	699	GAGAGGTTGCCTAATGCTGTTAAGCAAGAGTGCTGTTAGGCTTGCTTC	758	Db	1609	GACTGGGGAGTTGGAGCTTGTGACTGCTTGGGTTAGTGTAGATCTGAA	1668
Qy	1208	GAGTGCTGAGGGATGCTATGGATTGGATCCCTGATTATACGGTTAGGTT	1267	Qy	2288	CCAGCAGCTGGAACTGTTGGAGGGTGTCTTCCTAGGTCTAGAGCACCAA	2347
Db	759	GAGTGCTGAGGGATGCTATGGATCCCTGATTATACGGTTAGGTT	818	Db	1669	CCAGCAGCTGGAACTGTTGGAGGGTGTCTTCCTAGGTCTAGAGCACCAA	1728
Qy	1268	GTTCAGGAGCTTGTAGCTTACAGGTTAGTGTACTGCTTGGGTTAGGCTT	1327	Qy	2348	CAGATCAGAACTCGGGCTGAACCCAAGGATGTCGAAGAGCACAGTGTAGCTG	2407
Db	819	GTTCAGGAGCTTGTAGCTTACAGGTTAGTGTACTGCTTGGGTTAGGCTT	843	Db	1729	CAGATCAGAACTCGGGCTGAACCCAAGGATGTCGAAGAGCACAGTGTAGCTG	1788
Qy	1328	TTGGCTTATAAGAACTTCTGATTGATACTTGTATTGACTCTGTTAGGAGGA	1387	Qy	2408	TTGGTACATGTTAGGCGGTGAGGCTGGCCCTGGTCTTATGCAAGAGCTG	2467
Db	844	-----	851	Db	1789	TTGGTACATGTTAGGCGGTGAGGCTGGCCCTGGTCTTATGCAAGAGCTG	1848
Qy	1388	GGGCAAGTAGGCTGACCGGATTACGGCAAAATGATGAGCTGGAGAGATC	1447	Qy	2468	CCCTCTGAAACTTGAAACTAATGATGATGATGATGATGATGATGATGAG	2527
Db	852	GGGCAAGTAGGCTGACCGGATTACGGCAAAATGATGAGCTGGAGAGATC	911	Db	1849	CCCTCTGAAACTTGAAACTAATGATGATGATGATGATGATGATGAG	1908
Qy	1448	ACTCCGGTATCTCTGGACTCTGGCTTCTGGCTTGGATGATACGGCGAA	1507	Qy	2528	TTGATGAAACTACTGTGAAATGTCGCTTGGCTGATGATGATGATGAG	2587
Db	912	ACTCCGGTATCTCTGGACTCTGGCTTGGCTTGGATGATACGGCGAA	971	Db	1909	TTGATGAAACTACTGTGAAATGTCGCTTGGCTGATGATGATGATGAG	1968
Qy	1508	AGCTAAATGGTTAAGCGGTGCGGGATATTGTTGCTCTGGAGGGGAGGCA	1567	Qy	2588	TCCTAGCTGCTGGTGTGCAATTGTCGCTTGGCTGATGATGATGATGAG	2647
Db	972	AGCTAAATGGTTAAGCGGTGCGGGATATTGTTGCTCTGGAGGGGAGGCA	1031	Db	1969	TCCTAGCTGCTGGTGTGCAATTGTCGCTTGGCTGATGATGATGATGAG	2028
Qy	1568	TGAGCTCTGGTGTGGGTTGACCGCGTGAAGATTAGAATGGGTTTACGATG	1627	Qy	2648	AAAGCAGCTCACTTTCACCCAGGATATGGCTTCTCTGAGGATGTTGCTA	2707
Db	1032	TGAGCTCTGGTGTGGGTTGACCGCGTGAAGATTAGAATGGGTTTACGATG	1091	Db	2029	AAAGCAGCTCACTTTCACCCAGGATATGGCTTCTCTGAGGATGTTGCTA	2088
Qy	1628	ACGCTGTGAGGAGTACAGTTAGATACCTTTTAATTCCTAGGATATA	1687	Qy	2708	CCATAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2767
Db	1092	ACGCTGTGAGGAGTACAGTTAGATACCTTTTAATTCCTAGGATATA	1104	Db	2089	CCCA-----	2091
Qy	1688	ACTTAGTTCTCATTTAATGATGTTGTTGAGTGTGACCTGCTACCC	1747	Qy	2768	TTGGTACATGTTAGGAGCTTGTGAGGATGTTGAGGATGTTGAGGAT	2827
Db	1105	-----	1128	Db	2092	-----	2091
Qy	1748	CAAGCAATTTCAGCAGCTTGTGAGTGTGAGTGTGACCTGCTACCC	1807	Qy	2828	ACAGTGTGAGTAACTCACTAATGATGATGATGATGATGATGATGATG	2887
Db	1129	CAAGCAATTTCAGCAGCTTGTGAGTGTGAGTGTGACCTGCTACCC	1188	Db	2092	-----	2091
Qy	1808	AACCTTTATGGTAAAGGCCACCTTGTAGGAGTGTGAGTAACTGAC	1867	Qy	2888	CAGAATGTTAAATTCACTCAACATTCTGTTAGAATACGTTAGGATGAGATG	2947
Db	2092	-----	2091	Db	2092	-----	2091



QY	936	TTCAGCGACGACGCCCTTAATCACCGGAGACAGATCTCAAGCTGCCTGCGAAACTCT	420
QY	900	GTCATACTCCCGGCTAGAGAGAGTACATGAAGGCTCTCTGTGAGTGAGAGCTAC	959
Db	421	GTCATACCTCCGGCTTAGAGAGAGTACATGAAGGCTCTCTGTGAGTGAGAGCTAC	480
QY	950	AGTCATACCTGATGTTCTTGGATTAGGTAAATTGATTCAGTTGATGTCGATGATAAAGTTCT	1019
481	AGTCATACCTGATGTTCTTGGAT-		
Db	1020	CGTTTAATTCATGATAAAGGAGAACCTTATCTAGTGANGGTCCTGCGATAATAAAGTTCT	505
QY	506	-	
Db	1080	CTCTCTGTTATGCCAGAGGGTGTGAGACTGAGTAGTCTCCTGGGTTGGGGCTC	1139
Db	519	CTCTCTGTTATGCCAGAGGGTGTGAGACTGAGTAGTCTCCTGGGTTGGGGCTC	578
QY	1140	TGCTTAAGGAGAGGTGCTTAAGCTGAGAGTCTTGTGAGCTGAGTAGTCTCCTGGGTTGGGGCTC	1199
QY	579	TGCTTAAGGAGAGGTGCTTAAGCTGAGAGTCTTGTGAGCTGAGTAGTCTCCTGGGTTGGGGCTC	638
Db	1200	CGTTCTGATGCTGAGGATCTTGTGAGCTGAGTAGTCTCCTGGGTTGGGGCTC	1259
QY	639	CGTTCTGATGCTGAGGATCTTGTGAGCTGAGTAGTCTCCTGGGTTGGGGCTC	698
Db	1260	ATGAGTTGTTGAGGAAGCTTGAGCTTACAGGAGTTGAGCTGAGTAGTCTCCTGGGTTAGTGTATGGCTTG	1319
QY	699	ATGAGTTGTTGAGGAAGCTTGAGCTGAGTAGTCTCCTGGGTTAGTGTATGGCTTG	731
QY	1320	GACGGCGCTGGCTTATAAGAACTTCTGATGTTGATACTTTTATGAGCTGTGT	1379
Db	732	-	731
QY	1380	ACGAGGAGGAGGAGTACCTTGACCGATTAGTCGACAATTGATGAGACTTGG	1439
Db	732	ACGAGGAGGAGGAGTACCTTGACCGATTAGTCGACAATTGATGAGACTTGG	791
QY	1440	AAGAGATCACTCCGGTATGCTGGACTCTGGCTTACCGCTGGTGAAGTACG	1499
Db	792	AAGAGATCACTCCGGTATGCTGGACTCTGGCTTACCGCTGGTGAAGTACG	851
QY	1500	CTGGAAAAGACTTAAATGGTTAAGGGTGTGCGGAAATTGGGTGTGGAGGAG	1559
Db	852	CTGGAAAAGACTTAAATGGTTAAGGGTGTGCGGAAATTGGGTGTGGAGGAG	911
QY	1560	GTGAGCTCATGGCTTGTGGGGTTGACCGCTGAGGTTATGGATGAGCGGTT	1619
Db	912	GTGAGCTCATGGCTTGTGGGGTTGACCGCTGAGGTTATGGATGAGCGGTT	971
QY	1620	TACCATGACAGCTGCTGAGCGGTTACAGTTGATAACCTTTTAACTTGTAGC	1679
Db	972	TACCATGACAGCTGCTGAGC-	992
QY	1680	ATGATATACTTGTGTTCTCANTTTAATGATGTTGTGGTAGGTGATCTTGT	1739
Db	993	-	1008
QY	1740	AGCTTACCCAAAGCTATTCAGAGAGCTTACGAGCTTACGAGCTGACTGCT	1799
Db	1009	AGCTTACCCAAAGCTATTCAGAGAGCTTACGAGCTTACGAGCTGACTGCT	1068
QY	1800	TGTCGCTGAGCTTATGGTAAAGGCCACCTTACAGGTGCTGATAACATT	1859
Db	1069	TGTCGCTGAGCTTATGGTAAAGGCCACCTTACAGGTGCTGACTGCT	1128
QY	1860	CCAGCAACTTCAGGGCTAGGTATGGCTATGGAGTCTCGGATGTTGATGATC	1919
Db	1129	CCAGCAACTTCAGGGCTAGGTATGGCTATGGAGTCTCGGATGTTGATGATC	1188
QY	1920	ACGGATAATTGGAGATAGCTGGCTAGAGGAGACTCTGGCTCTGCTTATGG	1979
QY	1189	ACGGATAATTGGAGATAGCTGGCTAGAGGAGACTCTGGCTCTGCTTATGG	1248
QY	1980	CAAATGATGAACTCCGTTGTTGGCTTACAGTGAGGATCACATAAGGA	2039
Db	1249	CAAATGATGAACTCCGTTGTTGGCTTACAGTGAGGATCACATAAGGA	1308
QY	2040	TCCACCTATGTTGGAGTTGTTGGAGATTCATCGTGTGAACTGATGATCTCC	2099
Db	1309	TCCACCTATGTTGGAGTTGTTGGAGATTCATCGTGTGAACTGATGATCTCC	1368
QY	2100	TGGACTATGTCGAGTCTGGGTTGGAGATTCATCGTGTGAACTGATGAGA	2159
Db	1369	TGGACTATGTCGAGTCTGGGTTGGAGATTCATCGTGTGAACTGATGAGA	1428
QY	2160	CACCAAGATAAAATTRACTCGGACTACTATGATGATCTATGTTGAGT	2219
Db	1429	CACCAAGATAAAATTRACTCGGACTACTATGATGATCTATGTTGAGT	1488
QY	2220	CTTGGAAAGAGTGAGGAGTACTTCAGGGTCTCTTCTGCTCTGCAACTATGCAAG	2279
Db	1489	CTTGGAAAGAGTGAGGAGTACTTCAGGGTCTCTTCTGCTCTGCAACTATGCAAG	1548
QY	2280	GATGGAGCCGAGCATGAGAAAGCTGACTGCTGAGGACTATGATGATCTATGTTGAGT	2339
Db	1609	GATGGAGCCGAGCATGAGAAAGCTGACTGCTGAGGACTATGATGATCTATGTTGAGT	1668
QY	1549	TGACTGAGGAGCATGTCGAGTGTGCTATGGAGCTGAGGACTGAGAAGT	1608
Db	2340	CGGTATACAGATGAAACTCGGTGACCCAGATGTCAGAGACGCTGTTAGTGT	2339
QY	1609	CCGCTATACAGATGAAACTCGGTGACCCAGATGTCAGAGACGCTGTTAGTGT	1668
Db	1549	TGACTGAGGAGCATGTCGAGTGTGCTATGGAGCTGAGGACTGAGAAGT	1608
QY	2400	AGATCTGTTGTTACAACTGAGCGGTGAGCTGTTAGGAGCTGAGCAG	2459
Db	1669	AGATCTGTTGTTACAACTGAGCGGTGAGCTGAGCCTTATGAGAAGC	1728
QY	2460	TGTAAGACCCCTGTAACAACTTGAACTATGATGATCTGAGCTGGGCTGAG	2519
Db	1729	TGTAAGACCCCTGTAACAACTTGAACTATGATGATCTGAGCTGGGCTGAG	1788
QY	2520	GAGTAGCGTTGAGAACTACTTGAACTATGATGATCTGAGCTGGGCTGAG	2579
Db	1789	GAGTAGCGTTGAGAACTACTTGAACTATGATGATCTGAGCTGGGCTGAG	1848
QY	2580	TGTGAGATACTCTGCTGCTGGGGCAATTGGACTGATTCAGTTCTGTCAGCCAGAAGT	2639
Db	1849	TGTGAGATACTCTGCTGCTGGGGCAATTGGACTGATTCAGTTCTGTCAGCCAGAAGT	1908
QY	2640	TTCTTAAGAGCTCATCTTCTCAAGGAGGATATGGTCTCTCTATGGAACTCTGA	2699
Db	1909	TTCTTAAGAGCTCATCTTCTCAAGGAGGATATGGTCTCTCTATGGAACTCTGA	1968
QY	2700	TGTCGCTACATAGTGTGATTAATGATCCTATATCTGCAATTGCTCAAAAT	2759
Db	1969	TGTCGCTACATAGTGTGATTAATGATCCTATATCTGCAATTGCTCAAAAT	1979
QY	2760	ATGCTGTTGIGAGCTAAGAACATAGTCTCCACTTAATCATGTCAAAAGTGTAC	2819
Db	1980	-	1979
QY	2820	CAAGTTAACAGTGTGAGTAATTCTCACTTAATATGCTGCTGATTTTGATCAA	2879
Db	1980	-	1979
QY	2880	ACTGTTGAGACGAAATGTAATTCTCACTCTGTTGAGTACGGTGTGAGGATTA	2939
Db	1980	-	1979
QY	2940	GAGATTGCTTAGTGTGGCTTGTCCAACCTTCTCTGTTGATTTCTCTGATT	2999
Db	1980	-	1979
QY	3000	TAGGGTCACTGAGCTGAGCTGAGGAGACTCTGGCTCTGCTTATGG	3059
Db	1980	AGGGTCACTGAGCTGAGGAGACTCTGGCTCTGCTTATGG	2039

QY	3160	TGAGATAGTTACCAAGTGGAGGATAATCTAACATCAATTGTGAAACTG	3179	Arabidopsis Full-Length cDNA: Seki,M., Narusaka,M., Ishida,J., Hayashizaki,Y., and Shinozaki,K.
QY	2040	AGATATAGTATCCAGTGCGAGAATTAAGTCCTGCTGCTGATCACCGCA	2099	
Db	3120	TAGAAATTACCAAGAGGTGGAGGATAATCTAACATCAATTGTGAAACTG	3179	The Salk, Stanford, PGRC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the R AFL cDNA: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsian,V.W., Lee,J.M., Ouach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Paim,P., Shim,J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R., and Theologis,A.
Db	2115	-----	-----	-ASGTTTGTGATGG 2128
QY	3240	CGATGCTGAAGATTTGGACTGACAGAGCTGAACTGCGCAGCTGGTTTAT	3299	Yamada,K. (SSP/PGRC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGRC) contributed equally to this work as PIs.
Db	2129	CGATGCTGAAGATTTGGACTGACAGAGCTGAACTGCGCAGCTGGTTTAT	2188	Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.
QY	3300	GATTTACACTGTGAACATCTGTCAGCTGACAGTCTGAGATGAACTGG	3359	Location/Qualifiers
Db	2189	GATTTACACTGTGAACATCTGTCAGCTGACAGTCTGAGATGAACTGG	2248	
QY	3360	GCTCTGGGAGCACTTGAGGAGTCGCTCTCATGATGGTCATCCAGA	3419	
Db	2249	GCTCTGGGAGCACTTGAGGAGTCGCTCTCATGATGGTCATCCAGA	2308	
QY	3420	AACATGCTACTGATGTCAGAACCTACACACAGATGCGAAGTTCTGGTCAAGTC	3479	
Db	2309	AACATGCTACTGATGTCAGAACCTACACACAGATGCGAAGTTCTGGTCAAGTC	2368	
QY	3480	GGGTGAAATCACTGAGGCTGTCCTGCAATATACTCATATGAGATGTC	3539	
Db	2369	GGGTGAAATCACTGAGGCTGTCCTGCAATATACTCATATGAGATGTC	2428	
RESULT 6				
LOCUS	AY150490	AY150490	2436 bp mRNA linear PLN 23-SEP-2002	
DEFINITION	Arabidopsis thaliana unknown protein (AT5g42480)	mRNA, complete		
ACCESSION	AY150490			
VERSION	AY150490.1	GI:2329771		
KEYWORDS	Arabidopsis thaliana (thale cress)			
ORGANISM				
REFERENCE				
AUTHORS	Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsian,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Sarou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R., and Theologis,A.			
JOURNAL	Arabidopsis Open Reading Frame (ORF) Clones			
AUTHORS	(bases 1 to 2436)			
TITLE	Unpublished			
REFERENCE				
QY	1	Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsian,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Sarou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R., and Theologis,A.	47.6%; Score 1744; DB 8; Length 2436; Matches 2429; Conservative 79.2%; Pkd. No. 0; Mismatches 4; Indels 632; Gaps 5;	
QY	2	Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA	481 ATGGAGCTTGAGTCAGTCGGCATGGTCCTCCATTCCAATTATGCCGATTACCA	540
COMMENT	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of R AFL cDNAs (R AFL cDNA : 'RIKEN	1 ATGGAAGCTTGAGTCAGTCGGCATGGTCCTCCATTCCAATTATGCCGATTACCA	60	
QY	541 CGGGGAGGAACTCCAGTGGACACACTCTTCCGGTCTCCGGCAGCACCTCTTCAACTCTGCTCCGGCAGC	600		
Db	61 CCGGGAGGAACTCCAGTGGACACACTCTTCCGGTCTCCGGCAGCACCTCTTCAACTCTGCTCCGGCAGC	120		
QY	601 AAATGGGCCACCGCTCTCCACTCAATTCACTCCGATTCCTCTCC	660		

Db	121 AATGGGGCAGCCGCTCTCTCGACTCAA TTCACTCCATTCTCTCTCC 180
Qy	661 TTRGCCACCGCCACACCCSACACTCGTCTCTCGCACCATCTATGATCGRC 720
Db	181 TTRGCCACCGCCACACCCSACACTCGTCTCTCGCACCATCTATGATCGRC 240
Qy	721 GAAGCCCGTCCATCCCTGTATTCCTCAGGTTAGGAGCTCAACACATT 780
Db	241 GAAGCCCGTCCATCCCTGTATTCCTCAGGTTAGGAGCTCAACACATT 300
Qy	781 TTACCGATGGAATCAGAGAGGATCGAAGCTAGGGTTGAAACCCGGATTGGT 840
Db	301 TTACCGATGGAATCAGAGAGGATCGAAGCTAGGGTTGAAACCCGGATTGGT 360
Qy	841 TTRAGCGACGACCTTAATCAGCCGGAGACATCTTCAGTCAGCTGCAGAACCTG 900
Db	361 TTRAGCGACGACCTTAATCAGCCGGAGACATCTTCAGTCAGCTGCAGAACCTG 420
Qy	901 TCTAACTCTGCTCTAGAGAGAGTACATGAGGCTCTCTGATGAGAGCTACA 960
Db	421 TCTAACTCTGCTCTAGAGAGAGTACATGAGGCTCTCTGATGAGAGCTACA 480
Qy	961 GTCATCACTGATCTCTGGATAAGGTTATTGATTCGAAATAAAGTTCTC 1020
Db	481 GTCATCACTGATCTCTGGATAAGGTTATTGATTCGAAATAAAGTTCTC 504
Qy	1021 GRTTTAATTCTAGAATGGATAAGGAACTTTATGAGGAGTCTGGGG 1080
Db	505 -----ANGTTCTGGTC 518
Qy	1081 TCTCTGTTATTCAGAGGGCTGAGACTGAGATAGTTCTGGGGTGGAGGCT 1140
Db	519 TCTCTGTTATTCAGAGGGCTGAGACTGAGATAGTTCTGGGGTGGAGGCT 578
Qy	1141 GCTTAAGGAGGGTGCTAAGTGTAAAGCAGATGGTTAGTATGGCTTG 1200
Db	579 GCTTAAGGAGGGTGCTAAGTGTAAAGCAGATGGTTAGTATGGCTTG 638
Qy	1201 GRTCTCGATGTCGAGGGATCTATGGCATGGTGGATCACCCTGATTATCTG 1260
Db	639 GRTCTCGATGTCGAGGGATCTATGGCATGGTGGATCACCCTGATTATCTG 698
Qy	1261 TGTGTTGTTGAGGAACSTTGGAGTTTACGGTAGTTGACTCTGGTTAGTA 1320
Db	699 TGTGTTGTTGAGGAACSTTGGAGTTGACTCTGGTTAGTA 730
Qy	1321 ACAGAGCTTGTTTATAGAACCTTCITGATTTGAGCTCTGGCTA 1380
Db	731 -----A 731
Qy	1381 GGGGAGGAGGAGTACGGTCAACGGATTAGTCGACAAATTGAGAGCTTGA 1440
Db	732 GGAGGAGGAGCAGTAGCTACCGGATTAGTCACAAATTGAGCTTGA 791
Qy	1441 AGGATCACTCCCGTTAGCTGGAGCTACTGGCTTACCGCTGGTGAATTAGC 1500
Db	792 AGGATCACTCCCGTTAGCTGGAGCTACTGGCTTACCGCTGGTGAATTAGC 851
Qy	1501 TCGGAAAGACTTAATGGTTAGGGTGTGGGAATATTGGCTGGTGAAGGG 1560
Db	852 TCGGAAAGACTTAATGGTTAGGGTGTGGGAATATTGGCTGGTGAAGGG 911
Qy	1561 TCGGAAAGACTTAATGGTTAGGGTGTGGGAATATTGGCTGGTGAAGGG 1620
Db	912 TCGGAAAGACTTAATGGTTAGGGTGTGGGAATATTGGCTGGTGAAGGG 971
Qy	1621 ACCATGAGCAGCTGCTGGAGCTTACGGTGTGGGAATATTGGCTGGTGA 1680
Db	972 ACCATGAGCAGCTGCTGGTGA 991
Qy	1681 TGTATAACTTGTGGTTCTCATTTAATGATGTTGCTGGTGAAGGTTTTGTA 1740
Db	992 -----AGGTTGATCTTTGTA 1008
Qy	1741 GCTACCCAAAGCAATTCAGCAGAGCTTGAAGTTAGGAAGTGTCACTGCTC 1800
Db	1009 GCTACCCAAAGCAATTCAGCAGAGCTTGAAGTTAGGAAGTGTCACTGCTC 1068
Qy	1801 GTCGCTCAAGCTTTATGGTAAAGAAGCCACACCTTTACAGGATGCTGATGACATT 1860
Db	1069 GTCGCTCAAGCTTTATGGTAAAGAAGCCACACCTTTACAGGATGCTGATGACATT 1128
Qy	1861 CAGCAACTTCAGCAGGTTAGGTTAGCTAGGAGATTCTCTGGATGCTGATGACATT 1920
Db	1129 CAGCAACTTCAGCAGGTTAGGTTAGCTAGGAGATTCTCTGGATGCTGATGACATT 1188
Qy	1921 CGGATAATTGGAGATAGACTCGGCTGAGAACCTGTCAGCTGACTGCTTATAGG 1980
Db	1189 CGGATAATTGGAGATAGACTCGGCTGAGAACCTGTCAGCTGACTGCTTATAGG 1248
Qy	1981 AAAGTGTGAAAGCCGTTAGCTGGCTTGTGGCTTAGACAGTGGAGGATCACATAAGGAT 2040
Db	1249 AAAGTGTGAAAGCCGTTAGCTGGCTTGTGGCTTAGACAGTGGAGGATCACATAAGGAT 1308
Qy	2041 CCAGTATGGAGTTGGAGGTTAGGAAUTCAATCTGGATGACATGATGATCFCCT 2100
Db	1309 CCAGTATGGAGTTGGAGGTTAGGAAUTCAATCTGGATGACATGATGATCFCCT 1368
Qy	2101 GGACTATCCAATTGTTGGAACCTGGTGGAGGGTGTCTCTAGTTGAGAC 2160
Db	1369 GGACTATCCAATTGTTGGAACCTGGTGGAGGGTGTCTCTAGTTGAGAC 1428
Qy	2161 ACCAAGATAAAATAACTTCGGGACACTATGATGATGATCTAGGTTGAGCT 2220
Db	1429 ACCAAGATAAAATAACTTCGGGACACTATGATGATGATCTAGGTTGAGCT 1488
Qy	2221 TTGGAAGAGTGGAGGATGGTAGTCAGGGTTCTCTAGTGTGCTGCAATATGGAT 2280
Db	1489 TTGGAAGAGTGGAGGATGGTAGTCAGGGTTCTCTAGTGTGCTGCAATATGGAT 1548
Qy	2281 ATGGGACCCGACATCTGAGCTGGCTATGAGCTGGCTGAGAACGTTTCTC 2340
Db	1549 ATGGGACCCGACATCTGAGCTGGCTATGAGCTGGCTGAGAACGTTTCTC 1608
Qy	2341 CGCTATAGATAAGACTCGCTGAACCCAGGATGTGAGAGCAGGTTAGTGA 2400
Db	1609 CGCTATAGATAAGACTCGCTGAACCCAGGATGTGAGAGCAGGTTAGTGA 1668
Qy	2401 GATCCTGGPACACTGAGCCCGTGTAGCTGGCTGCTTATAGCAGAGCT 2460
Db	1669 GATCCTGGPACACTGAGCCCGTGTAGCTGGCTGCTTATAGCAGAGCT 1728
Qy	2461 GTPAGACCCCTGAGAAACTTGAACACTAATGATGAGCTGGCTGCTGAG 2520
Db	1729 GTPAGACCCCTGAGAAACTTGAACACTAATGATGAGCTGGCTGCTGAG 1788
Qy	2521 AGTACCTGGTGAACACTCTGGTGAATGTCGTTCTGATATGTAAGGAGGAGT 2580
Db	1789 ACTAGCTTGTGAGAACTTGAACACTCTGGTGAATGTCGTTCTGATATGTAAGGAGGAGT 1848
Qy	2581 GTGAGAGCTCTAGCTGCTGGTGTGGCAATTGGACTGATTCTACTGTCAGCAGAGT 2640
Db	1849 GTGAGAGCTCTAGCTGCTGGTGTGGCAATTGGACTGATTCTACTGTCAGCAGAGT 1908
Qy	2641 TTCTTAAAGCAGCTCATTTCAACGCAAGGATGTTCTCTGAAATCTGAT 2700
Db	1909 TTCTTAAAGCAGCTCATTTCAACGCAAGGATGTTCTCTGAAATCTGAT 1968
Qy	2701 GTCGCTCAAGCTTATGGTAAAGCAGGTTCTCTGAAATCTGAT 2760
Db	1969 GTCGCTCAAA 1978
Qy	2761 TGCTGTGTTGTGGACTAAGACATAGTCCCACCTAATCATGTCACAAAGTGTAC 2820
Db	1979 ----- 1978

QY	2821	AGGATAACAAGTGTCTGACTAATTCAATTATGCTGCTGAAATTGATCAA	2880	TITLE Submitted (20-Mar-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:basaki@niasaffrc.go.jp), URL: <a href="http://rgp.dna.affrc.go.jp/">http://rgp.dna.affrc.go.jp/</a> , Tel: +81-298-38-7441, Fax: +81-298-38-7468)
QY	2881	CCTGGACAGAAATGTAATTCTCTCACATTCTCTTGAGATAACGTTGATTAG	2940	COMMENT On Jan 21, 2004 this sequence version replaced gi:3812429. Genes were predicted from the integrated results of the following: GENSCAN ( <a href="http://CCR.081.mit.edu/GENSCAN.html">http://CCR.081.mit.edu/GENSCAN.html</a> ), FGENESH ( <a href="http://www.softberry.com/">http://www.softberry.com/</a> ), GeneMark.hmm ( <a href="http://opala.biolog.gatech.edu/GeneMark/">http://opala.biolog.gatech.edu/GeneMark/</a> ), GlimmerM ( <a href="http://www.tigr.org/tdb/glimmer/glimmer_form.html">http://www.tigr.org/tdb/glimmer/glimmer_form.html</a> ), RicEMM ( <a href="http://rgp.dna.affrc.go.jp/RicEMM/">http://rgp.dna.affrc.go.jp/RicEMM/</a> ), SplicePredictor ( <a href="http://bioinformatics.labstate.edu/cgi-bin/bp.cgi">http://bioinformatics.labstate.edu/cgi-bin/bp.cgi</a> ), SIm4 ( <a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a> ), gap2 ( <a href="http://www.tigr.org/software/glimmerm/">http://www.tigr.org/software/glimmerm/</a> ), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr ( <a href="http://ncbi.nlm.nih.gov/blast/db">http://ncbi.nlm.nih.gov/blast/db</a> ) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no.
QY	3001	ACGCTCAGTCAGCCCTGACCATCGAGAACCTCCAGAATGGATGCTGAGCTGCGA	3060	A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, "putative", and "-like protein". A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an "unknown" protein. A gene predicted by two or more gene prediction programs is classified as a hypothetical protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable "hypothetical" protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0575F10 clone has an overlap with O01020 (C02 (DDBJ: AP004078) clone at 5' end and with P0182F12 (DDBJ: AB005311) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <a href="http://rgp.dna.affrc.go.jp/GenomeSeq.html">http://rgp.dna.affrc.go.jp/GenomeSeq.html</a> .
QY	1980	AGGGTCACTGAGCTGAGCTGAGATGATGCTGAGCTGAGCTGAGCTGAGA	2039	FEATURES Source
QY	3061	GATATAGTAGTCAGTCAGTGCAGAGATAATCTCAATTCAATGTTGAACTG	3120	1. -150462 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="Taxon:39947" /chromosome="2" /clone="P0575F10" /complement.(4987..8112) /gene="P0575F10_1" /complement.[join(4987..5412,5520..5679,5764..5914, 6096..6247,6738..6859,7864..8112)] /gene="P0575F10_1" /note="Supported by full-length cDNA(s): AK10553" /complement.[join(5003..5412,5520..5679,5764..5914, 6096..6247,6738..6859,7864..8112)] /gene="P0575F10_1" /note="Supported by full-length cDNA(s): AK073059" /complement.[join(5315..5412,5520..5679,5764..5914, 6096..6247,6738..6859,7864..8112)] /gene="P0575F10_1" /note="contains EST(s): AU070160(E61534) contains full-length cDNA(s): AK10553,AK073059" /cdon_Start=1 /product="putative eukaryotic translation initiation factor 3, subunit 1 alpha, 35kDa" /protein_id=BAD07941_1" /db_xref="GI:41053010" /translation="MEDWDSDDFOPAVPSAKPEPLKSWADEDVDEDDVKESEWEER EKKPPPEVKPAKPSKGKAKGKQOASTSSVVDDEALDPLSKLQRQVLVED FSTTELFKGKGKSEKSLLDTPFKPSKSFABSYELANKLRYEKSPHYMLKINR LSMASLKGADADISSIAIAANEKIAKEBAAAGKKQGAKKKQHLHENDDDFPG
QY	1979	-----	1979	REFERENCE 1 Sasaki, T., Matsumoto, T. and Yamamoto, K. Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2, PAC clone: P0575F10. 2 Published Only in Database (2002) 3 (bases 1 to 150462) 4 Sasaki, T., Matsumoto, T. and Yamamoto, K.
REFERENCE	1	Oryza sativa (japonica cultivar-group) Bukayota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Gramineae; Oryzeae; Oryza;		AUTHORS
AUTHORS		Sasaki, T., Matsumoto, T. and Yamamoto, K.		JOURNAL
ORGANISM		Oryza sativa (japonica cultivar-group)		
VERSION	AP004885.3	GLI-1053009		
KEYWORDS				
SOURCE				
REFERENCE				
AUTHORS				
ORGANISM				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
ORGANISM				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				



QY	1203	TTCCTGATGTCGAGGGATGCTATAAGCATTTACTSGTATG	1262
Db	12606	ATGTGCGATCTATCAGGGATCTATGGCGAAGCCCTCCAGATGTAATTGGCTCTGG	12547
QY	1263	AGTTGCTGAGGAACTTGAGCTTACAGG-----TAG	1298
Db	12546	AGGCTCTGAGGAGSCTCTGAGCTTCTGAGGAACTGTTGCTTACTCCCTAC	12487
QY	1299	TTGAGCTGCTTGTGTTAATTGACAGGAGCGGTTAGG-----	1355
Db	12486	TTTTCATGGGCTTATATTGAAAGAGATRAACCTCTGAGATGGTT	12427
QY	1356	-----ATACTTGTGTTATGAGCTCTG-----	1376
Db	12426	GGGAACATGGCTATTTTATAAGATGTTGATTCTTGAAAGCAATAATAA	12367
QY	1377	-----TGTAGGAGAAGGAGCAGTAGCTTGACCGATTAGCGACAAA	1424
Db	12366	GGAACTCTCATACTGAGGAGATGGAGCACTGAGCTTACGCTACAGA	12307
QY	1425	TTGATGAGACTTGGAGAGATCACTCCGGTTATGCTCTGGAGTACTGGCTACCGC	1484
Db	12306	TTGATGAAACTCTCGAGGATTAACCTGCTGTTATGGAGCTCTCCCTCTCA	12247
QY	1485	TTGGGATGATTAACCTCGGCAAACACTAAATGTTAACGGGTTGCGGAAATTTGT	1544
Db	12246	TTGACACAGGATCATAAGAGGCCAAGAGGGCTCAAGGTCGAGAACATTG	12187
QY	1545	GGTCTTGTGAGGAGGTGGAGCATGGCTTGTGGGTTGCCCTGAGAGTTA	1604
Db	12186	GGAGGCTTSCAGAGGAGGATTGTTACCTTGAGGAGATTCTGAGGTTA	12127
QY	1605	TTAATGAGGGTTTACGATGAGCTGAGCTGAGG-----ATACAGTTAG	1655
Db	12126	TGAAGGAGGTTTGTGAGGATGACATCAATTGAGGATGTCAGTTGTTTAT	12067
QY	1656	ATACCTTTTTAAATTCTTGTACATGATAACTTTGAGCTTCTCAATTGTTAG	1715
Db	12066	GAAGGAGCTGCCTATTATATGCGAGGACATCTTGAAACAGCTGTTACCCCTTC	12007
QY	1716	TGT----GGGTAGGTGATCTTTGTAGCTACCCANGATAATTCCAGCAGTC	1770
Db	12006	TGTTGACCTCCACAGTGGATTCTTCAAAACACCGAATGAGCTTCTCTGATG	11947
QY	1771	TTTGAGTTTACGAGATGCACTGCTCTGGTCAGCTTATGTTAGAGGCCA	1830
Db	11946	TTGAAATTACATAGTACCTGACACATCTGCTCAGCAATTAAAGGCCA	11887
QY	1831	CACCTTTACAGGAGCTGATAAGGAATTCCGCACTCAGCAATTAAAGGCCA	1890
Db	11886	CAATTCATCATGATGCGGAGATCTTGTGAGCACTCCGAAATTCAAGGCT	11827
QY	1891	ATGGAGATCTCGGAGTGTATGATACAGGAAATTGGGAGTAGCTGGCTA	1950
Db	11826	CATTATGCTTATG-----ATATGAGATGGCTCATG-----	11791
QY	1951	GAAGGGAGCTGTGCACTGCTTACGGAAAGTGTGATGAGTGGCTGAGG	2010
Db	11790	GAAGGGCATCTGTCATCTGAGCTGGAGATGTAGGAGATGTGCTTGG	11731
QY	2011	TTAGCACTGAGGATCACATATGGAAATTCCGAACTATGGGACCTG	2070
Db	11730	ATTTGATGAGTCTACACATAGAGCTTACAGAGCCAAATTAGTTAATGAC	11671
QY	2071	TC---AAATGCTGAGACATGATATCTCCGAACTATGTTGAGACCTG	2127
Db	11670	TCTACGATCATGAGAGATGATCTTCCAGGCTGCAACCTTGGAGCTG	11611
QY	2128	TTGGAGGGGTGCTTCTAGGTGAGACACCAAGATAAAATTAACTCGGG	2187
Db	11610	CTTATCTTGTGAGGTTTCTAGGAGCAGAGATCTGGGCACTGCACTG	11551
QY	2188	GACTACTATGATGATCCTATGAGTTGAGTTAGCTGAAAGAGTGGAGGT	2247
Db	11550	GATTACTAGATGATGTCAGGAAATTAAAGCTTACCTGAGAAGGATGGGGGG	11491
QY	2248	TCTCTTGTAGTCGCTGCACTPATGGCAGGTTGGCTCTCGCT	2289
Db	11490	TCTCATTTGGCTGCTGCTGCTATGCAAACCTGGCTCAGCTGCACT	11431
QY	2290	GACCGTGTGAGGCTATGCTGAGGACTGCGAGAAGTTTCTCCCGTATACA	2349
Db	11430	GCTACTGTGAGTCAATGCTATTCAGGCTTCAGCTGAGATGAGACAG	11371
QY	2350	GATAGAAACTGGCTGACCCAGGATGCGAGAGACAGTGTGTTAGTGATCCTGTT	2409
Db	11370	TTGACAGGTTGAGCATGCAAATGAGATGCCCTGGGATATCTTGAGAAATT	11311
QY	2410	GGTACAACTGAGCGTGTGAGGCTGAGCCCTGAGGTTCTGCT	2451
Db	11310	GACCAGGAAATGCACTGCTCATGTCGAGAATGCCCTGAGATTATCTGCT	11251
QY	2452	GCAGAGGCTGAGACCCCTGAGAAACTATGAGATGCAATTGAGCTGG	2511
Db	11250	GGCSCACTGTGCTGACTGTGAGCTAATGGGCCAATATTTGCGCTGTAAGGGCC	11191
QY	2512	GCTCAGAGGAGTAGGGTGTGAGAAACTACTGTGAGAAATCTCGTGTGCTGATATGTAAG	2571
Db	11190	CTTGTGCTTATGAGTGTGAGCTGAGCTGAGTGTGCTGATATGTAAG	11131
QY	2572	GAGGAGAGCTGAGATCCCTGAGCTGCTGCTGAGCTTCACTGTCAGC	2631
Db	11130	GATGGTGTGAGTAATCTTCAATTGCTAATGAGTCAACACACTGGAAATTAC	11073
QY	2632	CAGAGTATTCTTAAGGAGTCATTTCAACGGAAAGATATGTTCTCTGAG	2691
Db	11072	CTCACGTTGTTACTGGTATCTGTTGGCTGAGTATGCTGCTGATCTAGCT	11013
QY	2692	GAATCTGATGTGAGATCCCTGAGCTGAGCTGCTGCTGAGCTTCAATATCTGCT	2751
Db	11012	TGCTGATATGTTAACGAGTACATCAATTATTTGCTCATACACAG	10953
QY	2752	CTCAAATATGCTGTTGTTGAGCTGAGCTGCTGCTGAGCTTCACTGTC	2811
Db	10952	CTACTATANITCTGTAATGATAATCATAGAGGAGTAACCTCTCTGATG	10893
QY	2812	AGTGTACCGAGATACAGTGTGCTGAGTAATTCTCAATTATGCTGCTG	2666
Db	10892	ATCCCTGACTATGAGGAGATCTGAGCTACATCATTTCTAGATGAGCTGAG	0833
QY	2867	ATTTGATCAAATGCTGTTGAGCTGAGCAATTAGTGTCCCCTTAATCATGTC	2926
Db	10832	AATATGTTGAGTAACTGTAGAGCAAACTTAATTCACTTCACACATTCTGTTAGAA	10773
QY	2927	TAAGTGTGAGTATGAGATGCTTAAACTTGACTTGTGATGATGAGCTG	2985
Db	10772	CATCATGCGAGGAGGTGATGTTAAACTTGACTTGTGATGATGAGCTG	10713
QY	2986	TTCTCTTGTGAGTTGGCTGAGCTGAGCTGAGCTTCCGAGATGGA	3045
Db	10712	CTCTTACCTATACCTGCTTCTATAGAGGATATGAGATAACTCTCTATAGGA	0653
QY	3046	TGCTAGGACTGAGGAATATGATCCAGTGGAGAGGAGTAAGCTCTGGCTTGG	3105
Db	10652	TATTTGGCTACGGATGACTCTTACAGTTGTTGAGTAACTCTCT	10602
QY	3106	GCCTGATCACCGCTAGAAATTGTTACAGGAGGAAATACCTCAATCA	3165
Db	10601	--TGTGAGTCAACACATGCTCAATGCACTGAGCAATTCTGAGAACTTCC	10545
QY	3166	TTCGCTGAAACCTGAGCATGTTAGTGTGCTGAGGAGTGTGAGCTGAGACTGCC	3282
Db	10544	CACTCGATGATATCCCGPTGATGAGGGCTTAATGCTGAGCTG	10485
QY	3223	TATAGGTTGATGGCGAATGCTGAGAATTGAGCTGAGGAGCTGAGACTGCC	3282

Db	10484	TATAGGTCTTGATGGCACATGCTAAAGTCGACTGACCGAGCAGGGATTGGC	10425	Db	121	TTCCTCCGACTCAATTCACTCCATTCCTCCCTCCCTGCACCGCCACCA	180
Qy	3283	AGCTTGGTTGGTTATGATTACACTTGTGAACTATCTGTGACAGTGTGACAGTCT	3342	Qy	677	CCACCGCCACTCTGTCCTCTCCACCATCTATGATGCCGACGCCAGTCCCCA	736
Db	10424	GTCATGGGGTCTGGGATATACATATCCCAGTGTGAGATGTATGACATCACTATCT	10365	Db	181	CCACCGCCACTCTGTCCTCTCCACCATCTATGATGCCGACGCCAGTCCCCA	240
Qy	3343	CGACGAGTGAACACCGTCTGGAGAACACTCTGGAGCTCTGGAGCTCTGGCTACTG	3402	Qy	737	TCCCATGTTCTACAGGTTAGAGCTAACAACTTCTTACCGATGAGATA	796
Db	10364	CCTAGATGTCCACAGGCACTCTGGAGCTCTGGAGCTCTGGCTACTG	10305	Db	241	TCCCATGTTCTACAGGTTAGAGCTAACAACTTCTTACCGATGAGATA	300
Qy	3403	ATTGGTTCATCGAGAACACATCTACTGATCTAGACCTACAGAACACTACACAGA	3462	Qy	797	GAAGAGCATTCGAGCTA	814
Db	10304	ATCTTACTGAGGCCAGAACATGATCATATGACACAAATACATACCCGTTATGAGA	10245	Db	301	GAAGAGCATTCGAGCTA	318
RESULT 8				RESULT 9			
AY198936		AY198936		I66494		I66494	
DEFINITION		Arabidopsis thaliana Sequence 341 bp flanking Ds3 end of Ds-Enhancertrap insertion from line ER7222.		DEFINITION		Sequence 14 from patent US 5670367.	
ACCESSION		AY198936		ACCESSION		166494	
VERSION		AY198936.1		VERSION		166494.1	
KEYWORDS		. Arabidopsis thaliana (thale cress)		SOURCE		Unknown.	
SOURCE		ORGANISM		ORGANISM		Unclassified.	
		Eukaryota; Viriplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		REFERENCE		1 (bases 1 to 7218)	
REFERENCE		1 (bases 1 to 341)		AUTHORS		Dorner, F., Scheiflinger, P. and Falkner, P. Gunter.	
AUTHORS		McCormick, P., Simorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R., McCombie, W.R. and Martienssen, R.A.		TITLE		Recombinant fowlpox virus	
TITLE		Arabidopsis genomic sequences flanking Ds enhancer and gene traps in transgenic lines		JOURNAL		Patent: US 5670367-A 14-23-SEP-1997;	
JOURNAL		Unpublished		FEATURES		Location/Qualifiers	
REFERENCE		2 (bases 1 to 341)				1. .7218	
AUTHORS		May, B.P., Simorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R., McCombie, W.R. and Martienssen, R.A.				/organism="unknown"	
TITLE		Direct Submission				/mol-type="unassigned DNA"	
JOURNAL		Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA					
COMMENT		http://genetrap.cshl.org/					
FEATURES		location/Qualifiers					
Source		1. .311					
		/organism="Arabidopsis thaliana"					
		/mol-type="genomic DNA"					
		/db_xref="Taxon:3702"					
		/clone="ER7222.Ds3.03.05.2002.Jw20.341"					
		/ecotype="landsberg"					
		/notes="transgenic line ER7222"					
ORIGIN							
Query Match	8.3%	Score 304.4; DB 8; Length 341;					
Best Local Similarity	99.4%	Pred. No. 3.3e-61; Mismatches 0; Indels 1; Gaps 1;					
Matches	316;						
Qy	497	ACCGCGGATTGGCTCTCCCGATTCGAGATACCCGGGACGCCAACAGC	556	Qy	690	CGTCTCTGCCACCATCTATTGATGTCGGACGCCAGTCCCCTGCGACGCCACCGCCACTCT	689
Db	2	ACCTCGGCACTGCTCCGCAATT-CATTAATGCCGATTACCCGGGACGCCAACAGC	60	Db	1258	YYY	1317
Qy	557	TCCGACGCGAACACCTCTACATCGCTCGCCACAAATGGCGACCGTC	616	Qy	750	CTTACGAGCTTATGGAGCTAACATTCCTAACCGATGCGATCGAAGAGCATTCGA	809
Db	61	TCCGACGCGAACACCTCTACATCGCTCGCCACAAATGGCGACCGTC	120	Db	1318	YY	1377
Qy	617	TTCTCTGACTCAATTCACTCGATCTCTCTCTCTCTCCACGCCACCA	676	Qy	810	AGTAGGGTTGAAACGGCCGCAATCGGTTAGCGAGCAGCGTTAATGAGCGCG	869
RESULT 10				Db	1378	YY	1437
AK083744				Qy	870	ACGATCTCAAGCTGCTT	889
LOCUS				Db	1438	CCAATCTCTATCTT	1457
DEFINITION							
ACCESSION							



I PYTSKBLVYKEYSKKLDCMDLNKUSDEKEETTKOISNNDLILIKEKKKKDUDTFI  
 SHNEKOKGKMLPFDGGDDNTDDEPDLRCHLDDNTDDEPDLRCHLDDNTD  
 YNDEDIFSLSMSITTDKRGKKNMCKKSBCVDURNKUSINYDNNNNKILSTIN  
 VODENDKLYKNEQKIQDIDYDNENDKLYKNEQKIQDIDYDNENDKLYKNEQK  
 LYNDENDKLYKNDQIKQDIDYDNENDKLYKNEQKIQDIDYDNENDKLYKNDQIKQ  
 REIDKGRKSNCITICKDMKLDBIKKKKDVENNNSSEDDVCFYELKF  
 ICSNNIEGNNTIDDTININVENVSNDTNYOSITFNQKYLNDLNLISQIFS"  
 repeat\_region 1704. .1740 /rpt\_type=tandem  
 repeat\_region 2309. .2446 /rpt\_type=tandem  
 repeat\_region 2310. .2443 /rpt\_type=tandem  
 repeat\_region 2313. .2449 /rpt\_type=tandem  
 repeat\_region 2718. .2893 /rpt\_type=tandem  
 repeat\_region 3065. .3311 /rpt\_type=tandem  
 repeat\_region 3776. .3805 /rpt\_type=tandem  
 repeat\_region 4039. .4094 /rpt\_type=tandem  
 repeat\_region 4260. .4295 /rpt\_type=tandem  
 repeat\_region 4402. .4439 /rpt\_type=tandem  
 gene /rpt\_type=tandem  
 repeat\_region /locus tag="PF14\_0239"  
 complement /join(>4519. .4588, 4895. .5573, 5806. .>6088)  
 /locus tag="PF14\_0239"  
 complement /join(>4519. .4588, 4895. .5573, 5806. .6088)  
 /codon\_start=1 /product="hypothetical protein"  
 /protein id="AN36852\_1"  
 /db\_xref="GI:23497307"  
 /translation "MBRNEELEKTEWEESEEINEVEELFLASSDNDLVKCESIRI  
 ILSLURKEFDIHNQKVKLTLLSRLNCRKLITCLECLNLSQLPKELTERMLI  
 BILFDMMKEEKEDCICDILTIMLNLTRCKEGVKVLDINDSNINKEKDNPFSF  
 FLANKLYFFPLIPKPIKINKNDSKYVSHVINTISIKESIVFFGNVAFANKCSDQI  
 LVNERRAIIPLIPKPIKINKNDSKYVSHVINTISIKESIVFFGNVAFANKCSDQI  
 NNEBEINONIHILLIMMKNSILVSCSVKSRIIILYLCNRDYSRDNIENVINKLEG  
 SNSREIAAT"  
 repeat\_region 4677. .4714 /rpt\_type=tandem  
 repeat\_region 5600. .5641 /rpt\_type=tandem  
 repeat\_region 5745. .5791 /rpt\_type=tandem  
 repeat\_region 6108. .6152 /rpt\_type=tandem  
 repeat\_region 6168. .6222 /rpt\_type=tandem  
 repeat\_region 6220. .6268 /rpt\_type=tandem  
 repeat\_region 6420. .6481 /rpt\_type=tandem  
 repeat\_region 6563. .6621 /rpt\_type=tandem  
 repeat\_region 6718. .6780 /rpt\_type=tandem  
 repeat\_region 6719. .6790 /rpt\_type=tandem  
 repeat\_region 7543. .7568 /rpt\_type=tandem  
 repeat\_region 7761. .7816 /rpt\_type=tandem  
 repeat\_region 7824. .7849 /rpt\_type=tandem  
 repeat\_region 7871. .7930 /rpt\_type=tandem

gene /rpt\_type=tandem  
 complement (<7953. .>8438)  
 /locus tag="PF14\_0240"  
 complement (<7933. .>8438)  
 /locus tag="PF14\_0240"  
 complement (>7953. .8438)  
 /locus tag="PF14\_0240"  
 /locus tag="PF14\_0240"  
 /codon\_start=1 /product="ribosomal protein L21, putative"  
 /protein id="AN36853\_1"  
 /db\_xref="GI:23497308"  
 /translation "MGKSGRKRSGRTRYKESKFRKHGECTANKYKLEKLYGDYDIV  
 CDS CDSTQOKGMPPVYIGKTGKPHTRGVGVLUVKVKGHLIEQKVCVRLEHVRSRC  
 NEDPLRKTKRNARLIEKAUKLNHEHNTKRITGEPRKPAAMIKVPPSKITIEPLPYEE  
 Y"  
 repeat\_region 8467. .8550 /rpt\_type=tandem  
 repeat\_region 8473. .8539 /rpt\_type=tandem  
 repeat\_region 8477. .8537 /rpt\_type=tandem  
 repeat\_region 8723. .8844 /rpt\_type=tandem  
 repeat\_region 8723. .8840 /rpt\_type=tandem  
 repeat\_region 8782. .8844 /rpt\_type=tandem  
 repeat\_region 9019. .9059 /rpt\_type=tandem  
 repeat\_region 9019. .9058 /rpt\_type=tandem  
 repeat\_region 9133. .9159 /rpt\_type=tandem  
 repeat\_region 9221. .9254 /rpt\_type=tandem  
 repeat\_region 9604. .9637 /rpt\_type=tandem  
 repeat\_region 9668. .9723 /rpt\_type=tandem  
 repeat\_region 10036. .10073 /rpt\_type=tandem  
 repeat\_region 10445. .10472 /rpt\_type=tandem  
 repeat\_region 10511. .10551 /rpt\_type=tandem  
 repeat\_region 10565. .>11080 /rpt\_type=tandem  
 complement (<10565. .>11080)  
 /locus tag="PF14\_0241"  
 complement (10565. .11080)  
 /locus tag="PF14\_0241"  
 /codon\_start=1 /product="basic transcription factor 3b, putative"  
 /protein id="AN36854\_1"  
 /db\_xref="GI:23497309"  
 /translation "MEKISBILAAARKLKEKMGSGNLROIGGKGSSARRKKVHKNSI  
 SNKINLILKKGASGYGVDETCYVRTGDTTEFLPKPKLCAQSNTVTVTGKNE  
 HKIDINLKFECKGKGNMDMLLEKIDPNKINLINKESGGDTPKREBEDQBVDP  
 LVNFPEVSIKE  
 repeat\_region 11149. .11458 /rpt\_type=tandem  
 repeat\_region 11514. .11584 /rpt\_type=tandem  
 repeat\_region 11514. .11579

Query Match 1.5%; Score 54; DB 3; Length 250029;  
 Best Local Similarity 46.7%; Pred. No. 0 16; Mismatches 195; Indels 0; Gaps 0;  
 Matches 171; Conservative

QY 11. TAAGGAGATACATTATAAGCAATTGCTGATTCACAGATTGCTGGCTA 70  
 Db 218336 TAATAGAATGATGATGTGATGTTATCTGTGATGATATTATTATTGAA 218395



Source	Medical Center Dr., Rockville
FEATURES	location/Qualit
source	/organism="Mus musculus"
misc_feature	/mol_type="genomic DNA"
misc_feature	/db_xref="Taxon:10090"
misc_feature	/chromosome="11"
misc_feature	/clone="RP23-103H9"
misc_feature	/clone_libr="RPCI-23"
misc_feature	14973..15082
misc_feature	/note="Sequence from uni-directional dgTP big dye terminator reads only."
misc_feature	15168..15188
misc_feature	/note="Sequence from uni-directional dgTP big dye terminator reads only."
misc_feature	133867..134035
misc_feature	/note="Sequence from uni-directional dgTP big dye terminator reads only."
misc_feature	186504..186516
misc_feature	/note="Sequence from uni-directional dgTP big dye terminator reads only."
misc_feature	190588..190730
misc_feature	/note="Sequence from uni-directional dgTP big dye terminator reads only."
GIN	terminator reads only."
query Match	est Local Similarity 56.5%; Score 53.8; DB 10; length 215524;
atches	100; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
SESSION	AE014829 AE01185
STON	AE014829.1 GI:23494865
WORDS	.
ORGANISM	Plasmodium falciparum 3D7
ERENCE	Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
UTHORS	1 (bases 1 to 250078)
JAMES	Gardiner, M.J., Hall, N., Rung, B., White, O., Berriman, M., Hyman, R.W., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyle, S., Chan, M.-S., Nene, V., Shallow, S.J., Suh, B., Peterson, J., Anguilo, S., Perteet, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M.A., Fairlamb, A.H., Praunholz, M.J., Ross, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.N., Mungall, C., Venner, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and Barrell, B.
TITLE	Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL	Nature 419 (6506), 498-511 (2002)
PUBLISHED	1256864
2	(bases 1 to 250078)
UTHORS	Gardner, M.J.
TITLE	Direct Submission
JOURNAL	Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
2	19703..19703
repeat_region	/rpt_type=tander
repeat_region	1..1842
repeat_region	/rpt_type=tander
repeat_region	1962..995
repeat_region	/rpt_type=tander
repeat_region	1165..1226
repeat_region	/rpt_type=tander
repeat_region	1165..1225
repeat_region	/rpt_type=tander
repeat_region	2096..2758
repeat_region	/rpt_type=tander
repeat_region	2261..2922
repeat_region	/rpt_type=tander
repeat_region	2990..3567
repeat_region	/rpt_type=tander
repeat_region	9425..9594
repeat_region	/rpt_type=tander
repeat_region	9426..9563
repeat_region	/rpt_type=tander
repeat_region	9556..9594
repeat_region	9840..9873
repeat_region	/rpt_type=tander
repeat_region	9880..9944
repeat_region	/rpt_type=tander
repeat_region	9886..9964
repeat_region	/rpt_type=tander
repeat_region	10025..1004
repeat_region	/rpt_type=tander
repeat_region	10079..10204
repeat_region	/rpt_type=tander
repeat_region	10080..10213
repeat_region	/rpt_type=tander
repeat_region	10225..10307
repeat_region	/rpt_type=tander
repeat_region	10301..10359
repeat_region	/rpt_type=tander
repeat_region	10383..10483
repeat_region	/rpt_type=tander
repeat_region	10396..10441
repeat_region	/rpt_type=tander
repeat_region	10693..10777
repeat_region	/rpt_type=tander
repeat_region	10715..10795
repeat_region	/rpt_type=tander
repeat_region	10870..11098
repeat_region	/rpt_type=tander
repeat_region	11196..11235
repeat_region	/rpt_type=tander
repeat_region	11254..11347
repeat_region	/rpt_type=tander
repeat_region	11383..12436
repeat_region	/rpt_type=tander
repeat_region	12702..12758
repeat_region	/rpt_type=tander
repeat_region	13355..13538
repeat_region	/rpt_type=tander
repeat_region	13528..13690
repeat_region	/rpt_type=tander
repeat_region	16935..20232
repeat_region	/rpt_type=tander
repeat_region	16935..16980
repeat_region	/rpt_type=tander
repeat_region	19528..19698
repeat_region	/rpt_type=tander

repeat\_region 26455. .26494 /rpt\_type-tandem  
repeat\_region 26541. .26574 /rpt\_type-tandem  
repeat\_region 19987. .20121 /rpt\_type-tandem  
repeat\_region 20045. .20265 /rpt\_type-tandem  
repeat\_region 20175. .20310 /rpt\_type-tandem  
repeat\_region 20402. .20538 /rpt\_type-tandem  
repeat\_region 20446. .20501 /rpt\_type-tandem  
repeat\_region 20622. .20823 /rpt\_type-tandem  
repeat\_region 20786. .20955 /rpt\_type-tandem  
repeat\_region 20935. .21030 /rpt\_type-tandem  
repeat\_region 21037. .21322 /rpt\_type-tandem  
repeat\_region 21269. .21467 /rpt\_type-tandem  
repeat\_region 21529. .21636 /rpt\_type-tandem  
repeat\_region 21590. .21594 /rpt\_type-tandem  
repeat\_region 21632. .21641 /rpt\_type-tandem  
repeat\_region 21632. .22001 /rpt\_type-tandem  
repeat\_region 21663. .22441 /rpt\_type-tandem  
repeat\_region 22727. .22829 /rpt\_type-tandem  
repeat\_region 22762. .22842 /rpt\_type-tandem  
repeat\_region 22875. .23014 /rpt\_type-tandem  
repeat\_region 22948. .23249 /rpt\_type-tandem  
repeat\_region 23670. .23959 /rpt\_type-tandem  
repeat\_region 23904. .24125 /rpt\_type-tandem  
repeat\_region 23989. .24150 /rpt\_type-tandem  
repeat\_region 24430. .25018 /rpt\_type-tandem  
repeat\_region 24527. .24663 /rpt\_type-tandem  
repeat\_region 24530. .24751 /rpt\_type-tandem  
repeat\_region 24617. .24655 /rpt\_type-tandem  
repeat\_region 25405. .25446 /rpt\_type-tandem  
repeat\_region 25464. .25567 /rpt\_type-tandem  
repeat\_region 25572. .25734 /rpt\_type-tandem  
repeat\_region 25763. .26236 /rpt\_type-tandem  
repeat\_region 26240. .26353 /rpt\_type-tandem  
repeat\_region 26382. .26411 /rpt\_type-tandem  
repeat\_region 15 RESULT 15  
repeat\_region AC116986/c AC116986 AC116986\_33321\_bp DNA linear INV 12-MAR-2003  
repeat\_region LOCUS Dicyostelium discoideum chromosome 2 map 2234041-256730 strain  
repeat\_region DEFINITION Dicyostelium discoideum chromosome 2 map 2234041-256730 strain  
repeat\_region AX4, complete sequence.  
repeat\_region AC116986 AC116425  
repeat\_region AC116986.2 GI:2882947  
repeat\_region HTG. Dicyostelium discoideum  
repeat\_region SOURCE Dicyostelium discoideum  
repeat\_region ORGANISM Mycetozoa; Dictyosteliida; Dicyostelium  
repeat\_region REFERENCE 1 (bases 1 to 33321)



Search completed: June 9, 2005, 21:46:56  
Job time : 15487 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## Om nucleic - nucleic search, using sw model

Run on: June 9, 2005, 16:58:33 ; Search time 2053 Seconds  
(without alignments)  
11072.555 Million cell updates/sec

Title: US-10-600-070-3  
Perfect score: 3667  
Sequence: 1 tgttctgtccatataaggagaat.....ctataacatataagggttacaa 3667

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgns\_6/ptodata/1/pubnra/US07\_PUBCOMB.seq: \*  
2: /cgns\_6/ptodata/1/pubnra/PC07\_NEW\_PUB.seq: \*  
3: /cgns\_6/ptodata/1/pubnra/US06\_NEW\_PUB.seq: \*  
4: /cgns\_6/ptodata/1/pubnra/US06\_PUBCOMB.seq: \*  
5: /cgns\_6/ptodata/1/pubnra/US07\_NEW\_PUB.seq: \*  
6: /cgns\_6/ptodata/1/pubnra/PC07US\_PUBCOMB.seq: \*  
7: /cgns\_6/ptodata/1/pubnra/US08\_PUBCOMB.seq: \*  
8: /cgns\_6/ptodata/1/pubnra/US09\_PUBCOMB.seq: \*  
9: /cgns\_6/ptodata/1/pubnra/US10\_PUBCOMB.seq: \*  
10: /cgns\_6/ptodata/1/pubnra/US11\_PUBCOMB.seq: \*  
11: /cgns\_6/ptodata/1/pubnra/US12\_PUBCOMB.seq: \*  
12: /cgns\_6/ptodata/1/pubnra/US13\_PUBCOMB.seq: \*  
13: /cgns\_6/ptodata/1/pubnra/US14\_PUBCOMB.seq: \*  
14: /cgns\_6/ptodata/1/pubnra/US15\_PUBCOMB.seq: \*  
15: /cgns\_6/ptodata/1/pubnra/US16\_PUBCOMB.seq: \*  
16: /cgns\_6/ptodata/1/pubnra/US17\_PUBCOMB.seq: \*  
17: /cgns\_6/ptodata/1/pubnra/US18\_PUBCOMB.seq: \*  
18: /cgns\_6/ptodata/1/pubnra/US19\_PUBCOMB.seq: \*  
19: /cgns\_6/ptodata/1/pubnra/US10\_PUBCOMB.seq: \*  
20: /cgns\_6/ptodata/1/pubnra/US11\_PUBCOMB.seq: \*  
21: /cgns\_6/ptodata/1/pubnra/US12\_PUBCOMB.seq: \*  
22: /cgns\_6/ptodata/1/pubnra/US13\_PUBCOMB.seq: \*  
23: /cgns\_6/ptodata/1/pubnra/US14\_PUBCOMB.seq: \*  
24: /cgns\_6/ptodata/1/pubnra/US15\_PUBCOMB.seq: \*  
25: /cgns\_6/ptodata/1/pubnra/US16\_PUBCOMB.seq: \*  
/cgns\_6/ptodata/1/pubnra/US17\_PUBCOMB.seq: \*

## ALIGNMENTS

RESULT 1  
US-10-600-070-3  
; Sequence 3, Application US/10600070  
; Publication No. US20040139500A1  
; GENERAL INFORMATION:  
; APPLICANT: Osteryoung, Katherine W.  
; APPLICANT: Vitha, Stanislav  
; APPLICANT: Koksharova, Olga A.  
; APPLICANT: Gao, Hong  
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
; TITLE OF INVENTION: Use  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; CURRENT FILING DATE: 2003-06-20  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 3  
; LENGTH: 3667  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; US-10-600-070-3

Query Match 100.0%; Score 3667; DB 19; Length 3667;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matchers 3667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qry 1 TGTTCTGTCCATATAAGGAGATAACATTATAAGGATTTGCTGATTCACAGATTG 60  
Db 1 TGTCTGTCCATATAAGGAGATAACATTATAAGGATTTGCTGATTCACAGATTG 60

Pred. No. 1 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	3667	100.0	3667	Sequence 3, Appli
2	3663.8	99.9	3667	Sequence 10, Appli
3	1968.4	53.7	2679	Sequence 227, Appli
4	1931	52.7	2637	Sequence 130, Appli
5	1724	47.0	2406	Sequence 1, Appli
6	1700.8	46.9	2406	Sequence 9, Appli
7	1717.6	46.8	2406	Sequence 128, Appli

Qy 61 CTTGGCTATAGATTCACTGGCTGTTGCTTACATTACATGCTATAATAGTTCTG 120  
 Db 61 CTTGGCTATAGATTCACTGGCTGTTGCTTACATTACATGCTATAATAGTTCTG 120  
 Qy 121 AATTACATTCAGTGGATGTTAAGAAAGAGGGATGTGGGTTTGCG 180  
 Db 121 AATTACATTCAGTGGATGTTAAGAAAGAGGGATGTGGGTTTGCG 180  
 Qy 181 TTAACTTAAGTAGTCAGAAGATAAGTCATGGTACTGCTCTATATGTGAA 240  
 Db 181 TTAACTTAAGTAGTCAGAAGATAAGTCATGGTACTGCTCTATATGTGAA 240  
 Qy 241 AAATCCCTGGCAGCTGTACATCATAGATGTTTATTTCAGTAGGGTCA 300  
 Db 241 AAATCCCTGGCAGCTGTACATCATAGATGTTTATTTCAGTAGGGTCA 300  
 Qy 301 AAATCCCTGGCAGCTGTACATCATAGATGTTTATTTCAGTAGGGTCA 360  
 Db 301 AAATCCCTGGCAGCTGTACATCATAGATGTTTATTTCAGTAGGGTCA 360  
 Qy 361 GTTTTGATTACTAACTCTCAAATCATAAACCTAGAGGAACTCTCCC 420  
 Db 361 GTTTTGATTACTAACTCTCAAATCATAAACCTAGAGGAACTCTCCC 420  
 Qy 421 CAGTCCTCAAAATGTAACGAAAGTTGGTAGCTGACCTAAAGACCTCC 480  
 Db 421 CAGTCCTCAAAATGTAACGAAAGTTGGTAGCTGACCTAAAGACCTCC 480  
 Qy 481 ATGGAGCTCTGTCAGTCGCGATGGTCCTCCCATTCATATGCCATTACCA 540  
 Db 481 ATGGAGCTCTGTCAGTCGCGATGGTCCTCCCATTCATATGCCATTACCA 540  
 Qy 541 CCCGGAACAAAGCTCGACCTAGCCACAAACCTCTACACTATCGTCGCCAGC 600  
 Db 541 CCCGGAACAAAGCTCGACCTAGCCACAAACCTCTACACTATCGTCGCCAGC 600  
 Qy 601 AAATGGCGCGCTCTCTCGCACTCTCGACTCAATTCCTCCCTCCCTCC 660  
 Db 601 AAATGGCGCGCTCTCTCGCACTCTCGACTCAATTCCTCCCTCCCTCC 660  
 Qy 661 TTAGCCACCGCCACCACCGCGACTCTCGCTCTCGCACACACTCTACATCG 720  
 Db 661 TTAGCCACCGCCACCACCGCGACTCTCGCTCTCGCACACACTCTACATCG 720  
 Qy 721 GAGGGCACTCCATCCATGATTCTACGGGTTAGGAGTCAAACACATTG 780  
 Db 721 GAGGGCACTCCATCCATGATTCTACGGGTTAGGAGTCAAACACATTG 780  
 Qy 781 TTAACCGATGGATCAGAGGATTCAGCTAGCTAGGGTTCGAACCGCCGATTGGT 840  
 Db 781 TTAACCGATGGATCAGAGGATTCAGCTAGGGTTCGAACCGCCGATTGGT 840  
 Qy 841 TTAGCGAGGACCTTTATCAGCCGGAGACATTCTCAAGCTCGCTGCGAACTCTG 900  
 Db 841 TTAGCGAGGACCTTTATCAGCCGGAGACATTCTCAAGCTCGCTGCGAACTCTG 900  
 Qy 901 TCTAATCTCGGTCTAGAGAGGATCATGGGCTCTCTGATGAGAGCTACA 960  
 Db 901 TCTAATCTCGGTCTAGAGAGGATCATGGGCTCTCTGATGAGAGCTACA 960  
 Qy 961 GRCATCACTGATGTTCTGGGATAAGCTTATTCGAAATAAATAGTTCTC 1020  
 Db 961 GRCATCACTGATGTTCTGGGATAAGCTTATTCGAAATAAATAGTTCTC 1020  
 Qy 1021 GRTTAATTCATGATGTTCTGGGATAAGCTTATTCGAAATAAATAGTTCTC 1080  
 Db 1021 GRTTAATTCATGATGTTCTGGGATAAGCTTATTCGAAATAAATAGTTCTC 1080  
 Qy 1081 TCTCTGTTATTCAGAGGGTGTGAGACTGAGATCTCTGGGTTGGAGGCTCT 1140  
 Db 1081 TCTCTGTTATTCAGAGGGTGTGAGACTGAGATCTCTGGGTTGGAGGCTCT 1140  
 Qy 1141 GCTTAAGGAGGAGGTGCTAAGTCGTTAAGCAGATGTTGCTGTTAGGCTGC 1200

Db 1141 GCTTAAGGAGGAGGTGCTAAGTCGTTAAGCAGATGTTGCTGCTG 1200  
 Qy 1201 GTTTCTGATGTCGAGGGAGCTGGCATGGATCCACCTGATTTTACTGCTT 1260  
 Db 1201 GTTTCTGATGTCGAGGGAGCTGGCATGGATCCACCTGATTTTACTGCTT 1260  
 Qy 1261 TGAGTTCTGAGGAAGCTTGAGCTTACAGGTTTACGGTTACTGCTCTGTAATG 1320  
 Db 1261 TGAGTTCTGAGGAAGCTTGAGCTTACAGGTTTACGGTTACTGCTCTGTAATG 1320  
 Qy 1321 ACCAGCGTTGGCTTATAAGAAGACTCTTGTGTTGATGACTTGTGTTGTA 1380  
 Db 1321 ACCAGCGTTGGCTTATAAGAAGACTCTTGTGTTGATGACTTGTGTTGTA 1380  
 Qy 1381 GGAGGAGGAGGAGTGGCTCTGAGCTGCTGAGCTGCTGAGCTTGGAA 1440  
 Db 1381 GGAGGAGGAGGAGTGGCTCTGAGCTGCTGAGCTGCTGAGCTTGGAA 1440  
 Qy 1441 AGAGATCCTCCGGTTATGCTTGGAGCTCTGGCTTACCGCTTGATGATTACG 1500  
 Db 1441 AGAGATCCTCCGGTTATGCTTGGAGCTCTGGCTTACCGCTTGATGATTACG 1500  
 Qy 1501 TCGAAAMAGCTAAATGTTAAGGGTTGACGGGTTACGTCACAAATGAGACCTCC 1560  
 Db 1501 TCGAAAMAGCTAAATGTTAAGGGTTGACGGGTTACGTCACAAATGAGACCTCC 1560  
 Qy 1561 TCGAAAMAGCTAAATGTTAAGGGTTGACGGGTTGCGGAAATTGTTGGCTCTGTTGAGGAGG 1620  
 Db 1561 TCGAAAMAGCTAAATGTTAAGGGTTGACGGGTTGCGGAAATTGTTGGCTCTGTTGAGGAGG 1620  
 Qy 1621 AGAACTGAGCTGCTGAGGATACAGTTAGTACCTTTTAATTTCTTAGCA 1680  
 Db 1621 AGAACTGAGCTGCTGAGGATACAGTTAGTACCTTTTAATTTCTTAGCA 1680  
 Qy 1681 TGTATTAACCTTGTGTTCTGTTTAACTTGTGTTGAGGTTTGTGTTGAGGTTTGTG 1740  
 Db 1681 TGTATTAACCTTGTGTTCTGTTTAACTTGTGTTGAGGTTTGTGTTGAGGTTTGTG 1740  
 Qy 1741 GCTAACCCAAAGCRATATTCAGCAGAGCTATTGAGTTTACGAGCTGACTTGTGTTGAGGTTTGTG 1800  
 Db 1741 GCTAACCCAAAGCRATATTCAGCAGAGCTATTGAGTTTACGAGCTGACTTGTGTTGAGGTTTGTG 1800  
 Qy 1801 GPGGCTCAGCTTATGTTAGGAGCCACCTTTACGGGATCTGCTGATGAACTTC 1860  
 Db 1801 GPGGCTCAGCTTATGTTAGGAGCCACCTTTACGGGATCTGCTGATGAACTTC 1860  
 Qy 1861 CAGCAACTTCTGGGAGGTTAGGATTAAGCTGATGGAGTTCTGGGATGTTGATGATA 1920  
 Db 1861 CAGCAACTTCTGGGAGGTTAGGATTAAGCTGATGGAGTTCTGGGATGTTGATGATA 1920  
 Qy 1921 CGGAAATTGGGAGGATCTGGGCTCTAGAAGGGGATCTGGGACTCTTATAGGC 1980  
 Db 1921 CGGAAATTGGGAGGATCTGGGCTCTAGAAGGGGATCTGGGACTCTTATAGGC 1980  
 Qy 1981 AAGTCTGATGCTCTGTTAGGAGGCTTACGGGTTAGACAGTGGAGGTTACAACTTATGGAA 2040  
 Db 1981 AAGTCTGATGCTCTGTTAGGAGGCTTACGGGTTAGACAGTGGAGGTTACAACTTATGGAA 2040  
 Qy 2041 CCAGTATGTTGGAGTTTGGGAGGTTCAATGTTGATGACATGATCTCCT 2100  
 Db 2041 CCAGTATGTTGGAGTTTGGGAGGTTCAATGTTGATGACATGATCTCCT 2100  
 Qy 2101 GGACTATGGGAACTGGGACTCTGGGAGGTTGCTCTGGGAGGTTGCTCTGGGAGG 2160  
 Db 2101 GGACTATGGGAACTGGGACTCTGGGAGGTTGCTCTGGGAGGTTGCTCTGGGAGG 2160  
 Qy 2161 ACCAAAGATAAAAATTAACTCGGGGACTCTGGGACTCTGGGACTCTGGGACTCTGGG 2220  
 Db 2161 ACCAAAGATAAAAATTAACTCGGGGACTCTGGGACTCTGGGACTCTGGGACTCTGGG 2220  
 Qy 2221 TTGAAAGGGAGGTTGGGAGGTTGGGAGGTTGGGAGGTTGGGAGGTTGGGAGG 2280

Db	2221	TTCGAAGAGTGGAGGAGTTCAAGGTTCTCTTGTGCTGCACTATGCCAAGG	2280	Qy	3361	CCTGGGGAAAGCACTCTGGAGGAGTCGTGCTACTGATTGGTCATCCAGAA	3420
Qy	2281	ATGGGACCGGAGGATGCGAACAGTGAAAGTGACTGCTACCTCAGGCACTSGAGAAGTTTCCCTCC	2340	Db	3361	CTCTGGGGAAAGCACTCTGGAGGAGTCGTGCTACTGATTGGTCATCCAGAA	3420
Db	2281	ATGGGAGCGGAGGATGCGAACAGTGAAAGTGACTGCTACCTCAGGCACTSGAGAAGTTTCCCTCC	2340	Qy	3421	ACATATGCTACTGATGTCAGGAACTTACACACAGATACAGAAGTTTCGGTCAGTCAG	3480
Qy	2341	CGCTATACAGATGAACTCGGTGAAACCAAGGATGTCAGAACAGACAGTGTGTTAGCTA	2400	Db	3421	ACATATGCTACTGATGTCAGGAACTTACACACAGATACAGAAGTTTCGGTCAGTCAG	3480
Db	2341	CGCTATACAGATGAACTCGGTGAAACCAAGGATGTCAGAACAGACAGTGTGTTAGCTA	2400	Qy	3481	GGTGAAAAATCACTGAAAGCTCTGTCATCATATACTCATATGTCAGTCATGTC	3540
Qy	2401	GATCTCTGTGTTACAATGTAGCGCGTGTGGTAGGCCTGTTATGCGAAGCT	2460	Db	3481	GGTGAAAAATCACTGAAAGCTCTGTCATCATATACTCATATGTCAGTCATGTC	3540
Db	2401	GATCTCTGTGTTACAATGTAGCGCGTGTGGTAGGCCTGTTATGCGAAGCT	2460	Qy	3541	GAGCTGGAGATCTCTGTCGTGAAATTCTCTCTAGTAGTTGTTAAATGAA	3600
Qy	2461	GTAGACCCCTCTGAAACCTTGAAACTATGCTGAACTGTCAGTCAG	2520	Db	3541	GAGCTGGAGATCTCTGTCGTGAAATTCTCTCTAGTAGTTGTTAAATGAA	3600
Db	2461	GTAGACCCCTCTGAAACCTTGAAACTATGCTGAACTGTCAGTCAG	2520	Qy	3601	ACACAAAATAATGAGCTCTGACACCCCTTCTCTGATCTAACTATAACATAGG	3660
Qy	2521	AGTAGCGGTGATGAACTACTGTGAAATGTCGCTGTCGTTAGT	2580	Db	3601	ACACAAAATAATGAGCTCTGACACCCCTTCTCTGATCTAACTATAACATAGG	3660
Db	2521	AGTAGCGGTGATGAACTACTGTGAAATGTCGCTGTCGTTAGT	2580	Qy	3661	GCTACAA 3667	
Qy	2581	GTCAGAGATCTCTACTCTGCTGCTGCGAATGCGCTGAAAGT	2640	Db	3661	GCTACAA 3667	
Db	2581	GTCAGAGATCTCTACTCTGCTGCTGCGAATGCGCTGAAAGT	2640	Qy	3361	CCTGGGGAAAGCACTCTGGAGGAGTCGTGCTACTGATTGGTCATCCAGAA	3420
Qy	2641	TTCTTAAGCGCTCATCTTCAAGCGAAAGTATGGTTCTPATGATCGAT	2700	Db	3361	CTCTGGGGAAAGCACTCTGGAGGAGTCGTGCTACTGATTGGTCATCCAGAA	3420
Db	2641	TTCTTAAGCGCTCATCTTCAAGCGAAAGTATGGTTCTPATGATCGAT	2700	Qy	3421	ACATATGCTACTGATGTCAGGAACTTACACACAGATACAGAAGTTTCGGTCAGTCAG	3480
Qy	2701	GTCGCTACCATAGTGTGTTAATGATCCAACTTCATATGCTGATGTCCTCAAATA	2760	Db	3421	ACATATGCTACTGATGTCAGGAACTTACACACAGATACAGAAGTTTCGGTCAGTCAG	3480
Db	2701	GTCGCTACCATAGTGTGTTAATGATCCAACTTCATATGCTGATGTCCTCAAATA	2760	Qy	3481	GGTGAAAAATCACTGAAAGCTCTGTCATCATATACTCATATGTCAGTCATGTC	3540
Qy	2761	TGCTGTGTTGTGAGCTGAAACATAGTGTCCAACTTAATACATGTCCAAACATGTCACC	2820	Db	3481	GGTGAAAAATCACTGAAAGCTCTGTCATCATATACTCATATGTCAGTCATGTC	3540
Db	2761	TGCTGTGTTGTGAGCTGAAACATAGTGTCCAACTTAATACATGTCCAAACATGTCACC	2820	Qy	3541	GAGCTGGAGATCTCTGTCGTGAAATTCTCTCTAGTAGTTGTTAAATGAA	3600
Qy	2821	AGCTTAACAGTGTGTTAATGATGCAATTTCATACATGTCATGTCCTCAA	2880	Db	3541	GAGCTGGAGATCTCTGTCGTGAAATTCTCTCTAGTAGTTGTTAAATGAA	3600
Db	2821	AGCTTAACAGTGTGTTAATGATGCAATTTCATACATGTCATGTCCTCAA	2880	Qy	3601	ACACAAAATAATGAGCTCTGACACCCCTTCTCTGATCTAACTATAACATAGG	3660
Qy	2881	CTCTAGACAGAAATGTAATTCTCAACATTCTCTGTTGAATACGGATTAG	2940	Db	3601	ACACAAAATAATGAGCTCTGACACCCCTTCTCTGATCTAACTATAACATAGG	3660
Db	2881	CTCTAGACAGAAATGTAATTCTCAACATTCTCTGTTGAATACGGATTAG	2940	Qy	3661	GCTACAA 3667	
Qy	2941	AGTTGCGCTTAGTGCGTGTGCAACITTTCTCTGATTTTCTCTGATT	3000	Db	3661	GCTACAA 3667	
Db	2941	AGTTGCGCTTAGTGCGTGTGCAACITTTCTCTGATTTTCTCTGATT	3000	Qy	3361	CCTGGGGAAAGCACTCTGGAGGAGTCGTGCTACTGATTGGTCATCCAGAA	3420
Qy	3001	AGGGTCACTCGAGGCTGAGGATTAGCGACTTCCAGAAATGATGCTGGCGAA	3060	Db	3361	CTCTGGGGAAAGCACTCTGGAGGAGTCGTGCTACTGATTGGTCATCCAGAA	3420
Db	3001	AGGGTCACTCGAGGCTGAGGATTAGCGACTTCCAGAAATGATGCTGGCGAA	3060	Qy	3421	ACATATGCTACTGATGTCAGGAACTTACACACAGATACAGAAGTTTCGGTCAGTCAG	3480
Qy	3061	GAATATAGTGTCACTGAGGAGATTAATCTCAATTCAATGCAATTGCTGTTGAGT	3120	Db	3421	ACATATGCTACTGATGTCAGGAACTTACACACAGATACAGAAGTTTCGGTCAGTCAG	3480
Db	3061	GAATATAGTGTCACTGAGGAGATTAATCTCAATTCAATGCAATTGCTGTTGAGT	3120	Qy	3481	GGTGAAAAATCACTGAAAGCTCTGTCATCATATACTCATATGTCAGTCATGTC	3540
Qy	3121	AGGAATAGTGTCACTGAGGAGATTAATCTCAATTCAATGCAATTGCTGTTGAGT	3180	Db	3481	GGTGAAAAATCACTGAAAGCTCTGTCATCATATACTCATATGTCAGTCATGTC	3540
Db	3121	AGGAATAGTGTCACTGAGGAGATTAATCTCAATTCAATGCAATTGCTGTTGAGT	3180	Qy	3541	GAGCTGGAGATCTCTGTCGTGAAATTCTCTCTAGTAGTTGTTAAATGAA	3600
Qy	3181	TGCAATGATGTTAGTGTGCTGCTGTTGAGT	3240	Db	3541	GAGCTGGAGATCTCTGTCGTGAAATTCTCTCTAGTAGTTGTTAAATGAA	3600
Db	3181	TGCAATGATGTTAGTGTGCTGCTGTTGAGT	3240	Qy	3601	ACACAAAATAATGAGCTCTGACACCCCTTCTCTGATCTAACTATAACATAGG	3660
Qy	3241	GATCTCTGTGTTACAATCTCTGTTGAGT	3300	Db	3601	ACACAAAATAATGAGCTCTGACACCCCTTCTCTGATCTAACTATAACATAGG	3660
Db	3241	GATCTCTGTGTTACAATCTCTGTTGAGT	3300	Qy	3661	GCTACAA 3667	
Qy	3301	ATTATACACTGTGAAACTATCTCTGAGT	3360	Db	3661	GCTACAA 3667	
Db	3301	ATTATACACTGTGAAACTATCTCTGAGT	3360	Qy	3361	CCTGGGGAAAGCACTCTGGAGGAGTCGTGCTACTGATTGGTCATCCAGAA	3420

Db	301	AAGTCCCTGGCAGCTGTCACATCATAGTTGTTTTTCAGTAGTTGGTCACT 360	Db	1381	GAGGGAGGAGAGAGTGCTGCACGGATTACGTGCCAAATTGATGAGACTTGG 1440
Qy	361	GTTTTGATTAACTTACTCTCAAAATCAAATCCATAAACCTAGGACCAA 420	Qy	1441	AGAGATCACCTCCCGTTATGTCITGGAGCTACTTGCTTCACGCCCTGGTGAATACG 1500
Db	361	GTTTTGATTAACTTACTCTCAAAATCAAATCCATAAACCTAGGACCAA 420	Db	1441	AGAGATCACCTCCCGTTATGTCITGGAGCTACTTGCTTCACGCCCTGGTGAATACG 1500
Qy	421	CACTCTCTCAATATGAAACAGAACAGTTTGAGTACCTAAAGCAGCTCC 480	Qy	1501	TSGAAAGACTTAATGCTTAAAGCGGATTGGGAATTGTTGGCTGTTGGAGAGG 1560
Db	421	CACTCTCTCAATATGAAACAGAACAGTTTGAGTACCTAAAGCAGCTCC 480	Db	1501	TGCAAAAGACTTAATGCTTAAAGCGGATTGGGAATTGTTGGCTGTTGGAGAGG 1560
Qy	481	ATCGAGCTCTGAGTCAGTCGCGATGGCTTCCATCCAAATTCCATAACCTAGAGCCAA 540	Qy	1561	TGAGCATTCAGCTTGTGGGGTTGACCGTGAGAGTTATGATGAGGGTTT 1620
Db	481	ATCGAGCTCTGAGTCAGTCGCGATGGCTTCCATCCAAATTCCATAACCTAGAGCCAA 540	Db	1561	TSGAGCATTCAGCTTGTGGGGTTGACCGTGAGAGTTATGATGAGGGTTT 1620
Qy	541	CCGGCGAGACAAGCTCGACCTAGCCACAACCTCTACACTATCTGCTCGCC 600	Qy	1621	ACCATGAGCTGCTGAGCAGSTATACTTTAAATCTTCTAGCA 1680
Db	541	CCGGCGAGACAAGCTCGACCTAGCCACAACCTCTACACTATCTGCTCGCC 600	Db	1621	ATGATGAGCTGCTGAGCAGSTATACTTTAAATCTTCTAGCA 1680
Qy	601	AATGGCGACGCTCTCTCGACCTCAATTCACTCGATTCCTCCCTCTCC 660	Qy	1681	TGATATACTTGTGTTCTATTTAGTTAGTTAGTTGTTGTTGTTGTTG 1740
Db	601	AATGGCGACGCTCTCTCGACCTCAATTCACTCGATTCCTCCCTCTCC 660	Db	1681	TGATATACTTGTGTTCTATTTAGTTAGTTAGTTGTTGTTGTTG 1740
Qy	661	TTCGCCACCGCCACCCACCCACCGCACTCTGTTCTGCCACCATCTATGTCGCC 720	Qy	1741	GCTACCCAGGATATTCACCGAGCTCAGCTTGTGTTCTCCACATTCTCTCC 1800
Db	661	TTCGCCACCGCCACCCACCCACCGCACTCTGTTCTGCCACCATCTATGTCGCC 720	Db	1741	GCTACCCAGGATATTCACCGAGCTCAGCTTGTGTTCTCCACATTCTCTCC 1800
Qy	721	GAGGCCACGCTCCATCCCATGTTCTACAGGTTAGAGCTCAACACATTC 780	Qy	1801	GTCGCTCAAGCTTTATGTTAGAGGCTTACCTTTAACTCTTCTAGCA 1860
Db	721	GAGGCCACGCTCCATCCCATGTTCTACAGGTTAGAGCTCAACACATTC 780	Db	1801	GTCGCTCAAGCTTTATGTTAGAGGCTTACCTTTAACTCTTCTAGCA 1860
Qy	781	TTACCGATGGATMCAGAGAGCATCTGAGCTAGCTGGTTGAAACCCGCAATTGGT 840	Qy	1861	CAGCAACTTCACAGGCTTAAGTGTGCTTGTGTTGTTGTTGTTGTTG 1920
Db	781	TTACCGATGGATMCAGAGAGCATCTGAGCTAGCTGGTTGAAACCCGCAATTGGT 840	Db	1861	CAGCAACTTCACAGGCTTAAGTGTGCTTGTGTTGTTGTTGTTGTTG 1920
Qy	841	TTAGCGCGAACCTTTATCAGCGGAGAACATTCTCAACGCTGCTGCGAACCTG 900	Qy	1921	CGAAATATTGGAGGATAPACTCTGCGCTCTAGAAAGGCGCATCTGSCACTCTTAPAGG 1980
Db	841	TTAGCGCGAACCTTTATCAGCGGAGAACATTCTCAACGCTGCTGCGAACCTG 900	Db	1921	CGAAATATTGGAGGATAPACTCTGCGCTCTAGAAAGGCGCATCTGSCACTCTTAPAGG 1980
Qy	901	TCMATCTCGGCTCTAGAGAGGCTACATGAGGGCTCTGTGTTGAGAGCTACA 960	Qy	1981	AAAGTGATGATGCCCTATGPGTTGGCTTAAAGGAGCTCA 2040
Db	901	TCMATCTCGGCTCTAGAGAGGCTACATGAGGGCTCTGTGTTGAGAGCTACA 960	Db	1981	AAAGTGATGATGCCCTATGPGTTGGCTTAAAGGAGCTCA 2040
Qy	961	GTCATCACTGATTCCTCTGGATAAGGTTGATTAATTCTGATGAGTTCTC 1020	Qy	2041	CCAGCTATTGCGAGTTGTTGGACAAATGGTGAATGATCTGCTCC 2100
Db	961	GTCATCACTGATTCCTCTGGATAAGGTTGATTAATTCTGATGAGTTCTC 1020	Db	2041	CCAGCTATTGCGAGTTGTTGGACAAATGGTGAATGATCTGCTCC 2100
Qy	1021	GTTTAATTCTGATGATGAGGAGGACTTATCTGAGGTTCTGGGGTGTCTGGGC 1080	Qy	2101	GGACTATGCAAACTGTCGGGTTGCGAGGTTGTTGCTTCTGTTCTGAGAC 2160
Db	1021	GTTTAATTCTGATGAGGAGGACTTATCTGAGGTTCTGGGGTGTCTGGGC 1080	Db	2101	GGACTATGCAAACTGTCGGGTTGCGAGGTTGTTGCTTCTGTTCTGAGAC 2160
Qy	1081	TCTCTGTTATGCAAGAGGGTGTGAGACTGAGACTCTCTGGGGTGTCTGGGC 1140	Qy	2161	ACCAAGATAAAAATTAACTCGGGGACTGTCGGGTTGCGAGGTTGTTGAGCTAC 2220
Db	1081	TCTCTGTTATGCAAGAGGGTGTGAGACTGAGACTCTCTGGGGTGTCTGGGC 1140	Db	2161	ACCAAGATAAAAATTAACTCGGGGACTGTCGGGTTGCGAGGTTGTTGAGCTAC 2220
Qy	1141	GCTTAAGGAGGAGTGCTTAAGCAAGAGTGTTTGTGTTGAGCTGGCTTC 1200	Qy	2221	TGGAAGAGGAGGAGGAGTGTGAGCTAGGGTTCTCTGCTGCTGAACTATGGAG 2280
Db	1141	GCTTAAGGAGGAGTGTGCTTAAGCAAGAGTGTTGAGCTGGCTTC 1200	Db	2221	TGGAAGAGGAGGAGTGTGAGCTAGGGTTCTCTGCTGCTGAACTATGGAG 2280
Qy	1201	GTTCTCTGATGTTGAGGGATCTATGCACTGGTTGAGTTGAGCTGGCTTC 1260	Qy	2281	ATGGAGCGGAGGATGTCAGCTAGGTTGAGCTGGCTTC 2340
Db	1201	GTTCTCTGATGTTGAGGGATCTATGCACTGGTTGAGTTGAGCTGGCTTC 1260	Db	2281	ATGGAGCGGAGGATGTCAGCTAGGTTGAGCTGGCTTC 2340
Qy	1261	TGTTCTGATGTTGAGGGATCTATGCACTGGTTGAGTTGAGCTGGCTTC 1320	Qy	2341	CCCTATCAGATGAGAACTCTGCGCTATGCACTGGCTTC 2340
Db	1261	TGTTCTGATGTTGAGGGATCTATGCACTGGTTGAGTTGAGCTGGCTTC 1320	Db	2341	CCCTATCAGATGAGAACTCTGCGCTATGCACTGGCTTC 2340
Qy	1321	ACGAGCGTGTGCTTATAAGACTCTGTTGAGCTGGCTTC 1380	Qy	2401	GATCGTTGGTACATGTTGAGCTGGCTTC 2460
Db	1321	ACGAGCGTGTGCTTATAAGACTCTGTTGAGCTGGCTTC 1380	Db	2401	GATCGTTGGTACATGTTGAGCTGGCTTC 2460
Qy	1381	GGGGAGGAGGAGTACCTGACCGGATTACGGTACAGCAATTGATGAGCTTGA 1440	Qy	2461	GTAGACCCCTGAAACTTGAACATGATTGCACTTGAGCTGGGCTCLAGAG 2520
Db	1381	GGGGAGGAGGAGTACCTGACCGGATTACGGTACAGCAATTGATGAGCTTGA 1440	Db	2461	GTAGACCCCTGAAACTTGAACATGATTGCACTTGAGCTGGGCTCLAGAG 2520

QY 2521 AGTAGCGTGTGAACTACTGTGAAATGTCGTTGCTGAAATGTTAAGGGCAAGT 2580  
Db 2521 AGTAGCGTGTGAAACTACTGTGAAATGTCGTTGCTGAAATGTTAAGGGCAAGT 2580  
QY 2581 GTGAAAGATCCTAGCTGGGGGCAATTGGACTGATTCAGTCAGCCAGAGPAT 2640  
Db 2581 GTGAAAGATCCTAGCTGGGGGCAATTGGACTGATTCAGTCAGCCAGAGPAT 2640  
QY 2641 TTCTCTAAAGCGCTCATCTTTCAGCGAAGGATAATGGTTCTCATGGAATCTGAT 2700  
Db 2641 TTCTCTAAAGCGCTCATCTTTCAGCGAAGGATAATGGTTCTCATGGAATCTGAT 2700  
QY 2641 TTCTCTAAAGCGCTCATCTTTCAGCGAAGGATAATGGTTCTCATGGAATCTGAT 2700  
Db 2701 GTGCGTACCATAGTGTAAATGTCGAATTCTCATATCTGCATGTCAAATA 2760  
QY 2701 GTGCGTACCATAGTGTAAATGTCGAATTCTCATATCTGCATGTCAAATA 2760  
Db 2761 TCTCTGTTTGTCAGTAAAGAACATAGTGTCCACTTAATACATGTCGCCAAAGTTGACC 2820  
QY 2761 TCTCTGTTTGTCAGTAAAGAACATAGTGTCCACTTAATACATGTCGCCAAAGTTGACC 2820  
Db 2761 TCTCTGTTTGTCAGTAAAGAACATAGTGTCCACTTAATACATGTCGCCAAAGTTGACC 2820  
QY 2821 AAGATTAACAAGTGTCTGAGTAAGACATAGTGTCCACTTAATACATGTCGCCAAAGTTGACC 2880  
Db 2821 AAGATTAACAAGTGTCTGAGTAAGACATAGTGTCCACTTAATACATGTCGCCAAAGTTGACC 2880  
QY 2881 CTGAGACAGAAATGTAATTCACTCACATTCTGTTAGATAACGTTAGATT 2940  
Db 2881 CTGAGACAGAAATGTAATTCACTCACATTCTGTTAGATAACGTTAGATT 2940  
QY 2941 AGATTCGCTTAGTGTGGCTTGTCCAACTTTCTTCCTGATTTTCTTCGATT 3000  
Db 2941 AGATTCGCTTAGTGTGGCTTGTCCAACTTTCTTCCTGATTTTCTTCGATT 3000  
QY 3001 AGGGTCACTGAGCTGAGGATTCAGACACTTCCAGAATGATGTCAGGACTG 3060  
Db 3001 AGGGTCACTGAGCTGAGGATTCAGACACTTCCAGAATGATGTCAGGACTG 3060  
QY 3061 GAATATAGTATCCAGTCAGAGATTAGTCCTGGCTTGGCTGATCACCGAT 3120  
Db 3061 GAATATAGTATCCAGTCAGAGATTAGTCCTGGCTTGGCTGATCACCGAT 3120  
QY 3121 AGAAATCTTACCAAGAGGTGAGGGAAATAATCTACATCAATCACTTGTGAAACGT 3180  
Db 3121 AGAAATCTTACCAAGAGGTGAGGGAAATAATCTACATCACTTGTGAAACGT 3180  
QY 3181 TGCACATGTTACCAAGAGGTGAGGGAAATAATCTACATCACTTGTGAAACGT 3240  
Db 3181 TGCACATGTTACCAAGAGGTGAGGGAAATAATCTACATCACTTGTGAAACGT 3240  
QY 3241 GATGCTGAGATTTGGACTGACAGGAGCTGAACTCGCAGCTTGGTTATG 3300  
Db 3241 GATGCTGAGATTTGGACTGACAGGAGCTGAACTCGCAGCTTGGTTATG 3300  
QY 3301 ATTATACACTGTGAAACTATCTGTGACAGTGTGACAGTCTCGCAGATGG 3360  
Db 3301 ATTATACACTGTGAAACTATCTGTGACAGTGTGACAGTCTCGCAGATGG 3360  
QY 3361 CTCCTGTTGAAAGCACTCTGGAGAGTCGCTCTCATGTTGCTCATCCAGAA 3420  
Db 3361 CTCCTGTTGAAAGCACTCTGGAGAGTCGCTCTCATGTTGCTCATCCAGAA 3420  
QY 3421 ACATGCTACTGTGAGACCTACACACAGATAGAAGTTCTGCTGCAAGTCAG 3480  
Db 3421 ACATGCTACTGTGAGACCTACACACAGATAGAAGTTCTGCTGCAAGTCAG 3480  
QY 3481 GGTGGAAATCACTGAAAGCTCTCTCTGCACTATACTCTGAGTGTGAACT 3540  
Db 3481 GGTGGAAATCACTGAAAGCTCTCTCTGCACTATACTCTGAGTGTGAACT 3540  
QY 3541 GAGCTTGGAGATTCCTCTGTTGAAATTCTCTCTGAACTGAACTGAACT 3600  
Db 3541 GAGCTTGGAGATTCCTCTGTTGAAATTCTCTCTGAACTGAACTGAACT 3600  
QY 3601 ACACAAAAATTAACGTCCTGGCACACCCTTCTCTGATCTAACATCAAGG 3660  
Db 3601 ACACAAAAATTAACGTCCTGGCACACCCTTCTCTGATCTAACATCAAGG 3660  
QY 3661 GCTACAA 3667  
Db 3661 GCTACAA 3667

RESULT 3  
US-10-739-930-27  
Sequence 227, Application US/10739930  
; GENERAL INFORMATION:  
; Publication No. US20040216190A1  
; APPLICANT: KOVALIC, David K.  
; TITLE OF INVENTION: NUCLEAR ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-21 (5337)B  
; CURRENT APPLICATION NUMBER: US10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO: 227  
; LENGTH: 2679  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER13643\_1  
; US-10-739-930-227

Query Match 53.7%; Score 1968.4; DB 20; Length 2679;  
Best local similarity 80.5%; Pred. No. 0; Mismatches 11; Indels 632; Gaps 5;  
Matches 2657; Conservative 2; GATTAACTTACTACTCAATCAAATCAAATCCATAAACCTAGACGACAAAGTC  
Db 428 TCAATATGAAACAGAACAAAGTTTGTAGCTAAAGACACTCCATGGAG 487  
Qy 62 TTCAATATGAAACAGAACAAAGTTTGTAGCTAAAGACACTCCATGGAG 121  
Db 488 CTCTGAGTCAGTCGCGCATGTGTCCTCCCAATTCAATTGCGCATTCACCGGA 547  
Qy 122 CTCTGAGTCAGTCGCGCATGTGTCCTCCCAATTCAATTGCGCATTCACCGGA 181  
Db 548 CGACAAAGCTCGACGAGTGCACACACCTCTACACTATCTGCTCCGCCAGCAATGG 607  
Qy 182 CGACAAAGCTCGACGAGTGCACACACCTCTACACTATCTGCTCCGCCAGCAATGG 241  
Db 608 CGACCGCTCTCTCGCGACTCAATTCACTCGATTCCTCCCTCCCTCTCGCCA 667  
Qy 242 CGACCGCTCTCTCGCGACTCAATTCACTCGATTCCTCCCTCTCGCCA 301  
Db 668 CGCGCACCAACACCGACACTCTCGCTCTCGCCACATTATGCGCCAGGCC 727  
Qy 302 CGCGCACCAACACCGACACTCTCGCTCTCGCCACATTATGCGCCAGGCC 361  
Db 728 AGTCGCCCATCCCATGTTCTACAGGTATTAGGAGCTCAACACATTTCTAACCG 787  
Qy 728 AGTCGCCCATCCCATGTTCTACAGGTATTAGGAGCTCAACACATTTCTAACCG 787  
Db 362 AGTCGCCCATCCCATGTTCTACAGGTATTAGGAGCTCAACACATTTCTAACCG 421  
Qy 788 ATGAAATCAGAGACCTCGAAGCTTGTGAAACCGCCGAAATCTGGTTCAAGG 847  
Db 422 ATGAAATCAGAGACCTCGAAGCTTGTGAAACCGCCGAAATCTGGTTCAAGG 481  
Qy 848 AGGAGCTTATCGCCGGAGCAGATTCTCAAGCTGCTGGAAACTCTGCTTAATC 907  
Db 482 AGGAGCTTATCGCCGGAGCAGATTCTCAAGCTGCTGGAAACTCTGCTTAATC 541  
Qy 908 CTGGCTCTAGAGAGAGTACATGAGGTTCTGATGATGAGGAGCTCAAGCTCA 967  
Db 542 CTGGCTCTAGAGAGAGTACATGAGGTTCTGATGATGAGGAGCTCAAGCTCA 601

QY	968	CCTATGTTCTTGGATAAGTAAATTGGATTTCGGAAATAAAGTTCTCTGTTAA	1027	Db	1430	TTGGAGTTGTTGGAGANTCAATGGATGACAATGATGATCCTGGACTPAT	1489
Db	602	CTATGTTCTTGGAT-----	618	QY	2108	GCAATTGTGGAAACCTGGTNGCAGGGTGCTTCCTAGGTCAGAGACCAAAG	2167
QY	1028	TTTCATGATTGATAAAGGAAGAACTTTACTAGTGAACGTCMGGGCTCTCTCTG	1087	Db	1490	GCAATTGTGGAAACCTGGTNGCAGGGTGCTTCCTAGGTCAGAGACCAAAG	1549
Db	619	-----AAGGTTCCGGTGTCTCTG	639	QY	2168	ATAAAATTAAACTCGGGACTACTATGATGCCTPATGCTGTTGAGTTCTGAA	2227
QY	1088	GTTCATGTTGAAACCTGGTNGCAGGGTGCTTCCTAGGTCAGAGACCAAAG	1147	Db	1550	ATAAAATTAAACTCGGGACTACTATGATGCCTPATGCTGTTGAGTTCTGAA	1609
Db	640	GTTCATGTTGAAACCTGGTNGCAGGGTGCTTCCTAGGTCAGAGACCAAAG	699	QY	2228	GAGTGGAGGTAGTCGGTTCTGGCTTCTGGCTCTGGCTGCACTATGGAGATGGG	2287
QY	1148	GAGAAGGTTGCTTAAGTCGTTAAGCAAGATGTCGTTCTGGCTTCTGGCTTC	1207	Db	1610	GACTGGAGGTAGTCGGTTCTGGCTTCTGGCTGCACTATGGAGATGGG	1669
Db	700	GAGAAGGTTGCTTAAGTCGTTAAGCAAGATGTCGTTCTGGCTTCTGGCTTC	759	QY	2288	CCAGCAGTCAGTGAAGCTAGTAGCTGCTTGAGCTGCAACTGGCTTCCGCTATA	2347
QY	1208	GAGTCTGAGGAGATGCTATGGATGTCGTTCTGGCTTCTGGCTTCTGGCTTC	1267	Db	1670	CCAGCAGTCAGTGAAGCTAGTAGCTGCTTGAGCTGCAAGAAGTTTCCCTCCGCTATA	1729
Db	760	GATGTCCTGGAGGATGCTATGGATGTCGTTCTGGCTTCTGGCTTCTGGCTTC	819	QY	2348	CAGATGAGACTGGCTGACCCAAAGATGTCGAAGAGAGTC	1387
QY	1268	GTTGAGGAGCTTGAAGCTTTACAGETAGTTGACTCTGCTTGAAATTACGAGCG	1327	Db	1730	CAGATGAGACTGGCTGACCCAAAGGATGTCGAAGAGAGTC	1789
Db	820	GTTGAGGAGCTTGAAGCTTTACAC-----	844	QY	2408	TTCGACATGTTAGGCCGATGTCGCTGAGCTGCAAGAAGCTTCCCTCCGCTATA	2467
QY	1328	TTCGCTTATAAGAACCTTCTGATTTGACTCTTGTGATTGGACTCTGTTGAGCT	845	Db	1790	TTCGACATGTTAGGCCGATGTCGCTGAGCTGCAAGAAGCTTCCGCTATA	1849
Db	845	-----AGGAGGAA	852	QY	2468	CTCTCTGAAACCTTGAAACTAATGATTATGCAATTGAGCTGCTGCTGCTG	2527
QY	1388	GGGCAACTGACCTGACCGGATTAGTCGACAAATTGATGAGACTTGGAGAGATC	1447	Db	1850	CTCTGAAACCTTGAAACTTGAATGATGAGCTGCTGCTGCTGCTGCTG	1909
Db	853	GGGCAACTGACCTGACCGGATTAGTCGACAAATTGATGAGACTTGGAGAGATC	912	QY	2528	TTCGATGAAACTACTGTTGAAAGTCGCTTGTGATGATGTTAAGGAGGCAAGTGGAGA	2587
QY	1448	ACTCCCGTTATGCTCTGGACTACTGGCTTACCGCTTGGATGATTACGGCGAA	1507	Db	1910	TTCGATGAAACTACTGTTGAAATGCTGCTGCTGCTGCTGCTGCTG	1969
Db	913	ACTCCCGTTATGCTCTGGACTACTGGCTTACCGCTTGGATGATTACGGCGAA	972	QY	2588	TCTTAGCTGCTGCTGCAACTTGAGCTTAACTGATGATGCTGCTGCTGCTG	2647
QY	1508	AGCTAAATGGTTAAGGGTTGGGATATTGGCTCTGGCTGAGGACCA 1567	973	Db	1970	TCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2029
Db	973	AGCTAAATGGTTAAGGGTTGGGATATTGGCTCTGGCTGAGGACCA 1032	QY	2648	AAAGCAGTCATCTTCACCCAAAGGATATGGCTCTGGCTGCTGCTGCTG	2707	
QY	1568	TCTGGCTCTGGTGGGGTTGACCCGTTGAGAAGATTATGAACTGGGTTTACGATG	1627	Db	2030	AAAGCAGTCATCTTCACCCAAAGGATATGGCTCTGGCTGCTGCTGCTG	2089
Db	1033	TGAGCTCTGGTGGGGTTGACCCGTTGAGAAGATTATGAACTGGGTTTACGATG	1092	QY	2708	CCATAGCTGATGTTAAATGTCGAAATTTCATPATATCTGCTGCTGCTGCTG	2767
QY	1628	ACGGCTCTGGTGGGGTTGACCCGTTGAGAAGATTATGAACTGGGTTTACGATG	1687	Db	2090	CCAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2092
Db	1093	ACGGCTCTGGATC-----	1105	QY	2768	TTTGAGCTGAAACATGATGTCCTTAATACATGCCAAAGCTGTCACAA	2827
QY	1688	ACTTAGGTTCTCATTTAATGATGTTGTTGAGCTTGGAGCTTGGAGCTACC	1747	Db	2093	-----	2092
Db	1106	-----AGGTGATCTTTGGCTACC	1129	QY	2828	ACAAGTGTGATGAAATTCACTAATTATGCTGCTGATCAAATGCTG	2887
QY	1748	CTAGCAATATTGGAGGAGCTTGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	1807	Db	2093	-----	2092
Db	1130	CTAGCAATATTGGAGGAGCTTGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	1189	QY	2888	CAGAAATGTAATTCACTTCACAACTTCTGCTGATGTTGAGATGAGATG	2947
QY	1808	AAGCTTTATGGTGGAGCTTGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	1867	Db	2093	-----	2092
Db	1190	AAGCTTTATGGTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	1249	QY	2948	CTAGTGTGCTTGTCAACTTCTGCTGATGTTTCTGCTGATGTTGAGCT	3007
QY	1868	TTGAGCTGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	1927	Db	2093	-----TAGGTCA	2100
Db	1250	TTGAGCTGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	1309	QY	3008	GTCAGAGCTGACGATTGAGACCACTCCAGATGGATGCTAGAGCTGAGA	3067
QY	1928	ATGGGAGATGAGCTTGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	1987	Db	2101	GTCAGAGCTGACGATTGAGACCACTCCAGATGGATGCTAGAGCTGAGA	2160
Db	1310	ATGGGAGATGAGCTTGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	1369	QY	3068	GTATCGAGTCAGGAGATGAGACCACTCCAGATGGATGCTAGAGCTGAGA	3127
QY	1988	ATGAACTGGCTTGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	2047	Db	2161	GTATCGAGTCAGGAGATGAGACCACTCCAGATGGATGCTAGAGCTGAGA	2220
Db	1370	ATGAACTGGCTTGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCT	1429	QY	3128	TTACCGAGGTGAGGAAATACTCAATGATGAGCAATGATGATCTCCCTGGACTAT	2107











RESULT 7  
 US-10-600-070-128 Application US/10600070  
 ; Sequence 128, Application No. US10600070  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Everyoung, Katherine W.  
 ; APPLICANT: Vlitha, Stanislav  
 ; APPLICANT: Koksharova, Olga A.  
 ; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
 ; TITLE OR INVENTION: Use  
 ; FILE REFERENCE: MSU-08153  
 ; CURRENT APPLICATION NUMBER: US/10/600, 070  
 ; CURRENT FILING DATE: 2003-06-20  
 ; NUMBER OF SEQ ID NOS: 206  
 ; SOFTWARE: Patentin version 3.2  
 ; SEQ ID NO: 128  
 ; LENGTH: 2406  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-10-600-070-128  
 Query Match 46.8%; Score 1717.6; DB 19; Length 2406;  
 Best Local Similarity 79.1%; Pred. No. 0; Mismatches 4; Indels 632; Gaps 5;  
 Matches 2402; Conservative 0;  
 Qy 481 ATCGAAGCTCTGJTCAGTCGGCATGGTCTCCCATCCATTATGCCATTACCA 540  
 Db 1 ATCGAAGCTCTGJTCAGTCGGCATGGTCTCCCATCCATTACCA 60  
 Qy 541 CCGCGACGACAAGCTCGACGACTAGCCACAAACCTCTACAACTATGTGCCGCG 600  
 Db 61 CGCGCGACGACAAGCTCGACGACTAGCCACAAACCTCTACAACTATGTGCCGCG 120  
 Qy 601 AAATGGCGCGCTCTCTCGACGCTCAATTCACTTCACTCGCAATCTCTCTC 660  
 Db 121 AAATGGCGCGCTCTCTCGACGCTCAATTCACTTCACTCGCAATCTCTCTC 180  
 Qy 661 TTGCCAACCGCCACACCGCCACCTCGTCTCTGCCACATCTATGATGCC 720  
 Db 181 TTGCCAACCGCCACACCGCCACCTCGTCTCTGCCACATCTATGATGCC 240  
 Qy 721 GAGGCCAGTCCATCCATGATGTTGATGTTGATGTTGATGTTGATGCC 780  
 Db 241 GAGGCCAGTCCATCCATGATGTTGATGTTGATGTTGATGTTGATGCC 300  
 Qy 781 TTAACCGATGAGATCAGAGAGATCGAACGCTTGAAGGGCTCGCAACTCTG 840  
 Db 301 TTACCGATGAGATCAGAGAGATCGAACGCTTGAAGGGCTCGCAACTCTG 360  
 Qy 841 TTACCGATGAGATCAGAGAGATCGAACGCTTGAAGGGCTCGCAACTCTG 900  
 Db 361 TTACCGATGAGATCAGAGAGATCGAACGCTTGAAGGGCTCGCAACTCTG 420  
 Qy 901 TCTATCTCGGTCTAGAGAGATCGAACGCTTGAAGGGCTCGCAACTCTG 960  
 Db 421 TCTATCTCGGTCTAGAGAGATCGAACGCTTGAAGGGCTCGCAACTCTG 480  
 Qy 961 GTCATCACTGATGTTGGGATAAGGTTGATGTTGCAATATAAGTTCTC 1020

Db 481 GTCACTACTGAGTCTCTGGCAT----- 504  
 Qy 1021 GTTAAATTCTATGANTGGATAAGGAGGAACTTTATCTAGTGAGTGCTCTGGGC 1080  
 Db 505 -----  
 Qy 1081 TCTCTGTTATGCAAGGGGGTGTGAGCTGAGATGTTCTCTGGGTGGGGCTCT 1140  
 Db 519 TCTCTGTTATGCAAGGGGGTGTGAGCTGAGATGTTCTCTGGGTGGGGCTCT 578  
 Qy 1141 GCTTAACGAAGCGTGCCTAAGCTGTTAAGCAGAGTGTGGTTAGTTATGCCGCTTG 1200  
 Db 579 GCTTAACGAAGCGTGCCTAAGCTGTTAAGCAGAGTGTGGTTAGTTATGCCGCTTG 638  
 Qy 1201 GRTTCTCGATGTTCTGGGGGCTATGGCATGGATCCACTGATTTACTSGTTA 1260  
 Db 639 GRTTCTCGATGTTCTGGGGGCTATGGCATGGATCCACTGATTTACTSGTTA 698  
 Qy 1261 TGAGTTGTTGAGAACCTTGAGCTTACAGGTTAGCTGACTTGCTTGTAATTG 1320  
 Db 699 TAGTTGTTGAGAACCTTGAGCTTACAGGTTAGCTGACTTGCTTGTAATTG 730  
 Qy 1321 AGAGCGTTGGTTATAAGACTTTGTTGATTTGAGTTGAGCTTGTGTAATTG 1380  
 Db 731 -----  
 Qy 1441 AGAGATCACTCGGTTATGTTGAGCTCTGGCTTACCGCTTGGCTTGATGATGCG 1500  
 Db 792 AGAGATCACTCGGTTATGTTGAGCTCTGGCTTACCGCTTGGCTTGATGATGCG 851  
 Qy 912 TCGGAAAGACTAAATGTTAAGCGTTGGGAATATTGTTGTTGCTTGTGAGGAG 1560  
 Db 852 TCGGAAAGACTAAATGTTAAGCGTTGGGAATATTGTTGCTTGTGAGGAG 911  
 Qy 1561 TGGAGATCACTGCTCTGTTGGGGTTGACCGGTGAGAAGTTATGATGAGGCGTTT 1620  
 Db 912 TGGAGATCACTGCTCTGTTGGGGTTGACCGGTGAGAAGTTATGATGAGGCGTTT 971  
 Qy 1621 AGGAATGACGCTGCTGAGCAGGTTACAGTTAGATAACCTTTAAATCTTAGCA 1680  
 Db 972 AGGAATGACGCTGCTGAGC----- 991  
 Qy 1681 TGATATRACTTGTGTTCTCTTATGATGTTGTTGCTGAGTTGATCTTTGTA 1740  
 Db 992 -----  
 Qy 1741 GCTAACCCAAAGCAATATTCAGCAGACTCATTGAAAGTTGAGGTGACTTGCTCT 1800  
 Db 1009 GTCATCCCAAGCAATATTCAGCAGACTCATTGAAAGTTGACTTGCTCT 1068  
 Qy 1801 GPGGCTCAAGCTTATGTTGAGAACGCCACCTTACAGGATGCTGATGACATT 1860  
 Db 1069 GTCATCCCAAGCAATATTCAGCAGACTCATTGAAAGTTGACTTGCTCT 1128  
 Qy 1861 CGCAACTTCAAGCGTTAGCTGAGCTTGGAGTTCTGGATGTTGTTATGATCA 1920  
 Db 1129 CGCAACTTCAAGCGTTAGCTGAGCTTGGAGTTCTGGATGTTGCTCT 1188  
 Qy 1921 CGGAAATTGAGGAGATGACTCTGGCTTGAAGGGGACTCTGAGCTTATAGGC 1980  
 Db 1189 CGGAAATTGAGGAGATGACTCTGGCTTGAAGGGGACTCTGAGCTTATAGGC 1248  
 Qy 1981 AAAGTGTGATGAGCTGCTTGGGGCTTAGACGCTGAGGATTCACAAATAGGAT 2040  
 Db 1249 AAAGTGTGATGAGCTGCTTGGGGCTTAGACGCTGAGGATTCACAAATAGGAT 1308  
 Qy 2041 CGGAAATTGAGGAGATGACTCTGGCTTGAAGGGGACTCTGAGCTTATAGGC 2100



```

Db 414 AGTCTCGGTTGGGCTGTCAACCATGAACTTACAG 370
Qy 3150 TCTACAAATCAAATCAATTGTGCAAAGACTGTGATGTTCTGGCTTGTT 3209
Db 369 ----- 370
Qy 3210 TCAATTGTTATTAGTTGGTGGGGAATGCGAAGTATTGAGCTGACAGCA 3269
Db 369 ----- -AGGTTTGGATGGGGGAGATGCGAAGTGGACTGACAGAAC 326
Qy 3270 GCTGAAACTGCGAGCTGGTGGTTATGATACTGTGAACTATGTGAC 3329
Db 325 GCTGAACTGCGAGCTGGTGGTTATGATACTGTGAACTATGTGAC 266
Qy 3330 AGTGTGACAGTCAGCAGATGAAACCGTGCCTGGGAGCACTTGGAGCT 3389
Db 265 AGTGTGACAGTCAGCAGATGAAACCGTGCCTGGGAGCACTTGGAGCT 206
Qy 3450 ACGAGATCGAAAGTTCTGGTCAAGTCAGGGCTGTCT 3509
Db 145 ACAGATAGGAGTTTGGTCAAGTCAGGGCTGTCT 86
Qy 3510 GCATCATATACATATAGGAGCTGGAGATTCTTGTCTCTCTAA 3569
Db 85 GCATCATATACATATAGGAGCTGGAGATTCTTGTCTCTCTAA 26
Qy 3570 ATTCCTCTTAAAGTAGTTAT 3594
Db 25 ATTCCTCTTAAAGTAGTTAT 1

RESULT 9
US-10-600-070-184
; Sequence 184 Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteryoung, Katherine W.
; APPLICANT: Vitha, Stanislav
; APPLICANT: Roksharova, Olga A.
; APPLICANT: Gao, Hongo
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 184
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Prunus persica
; FEATURE: misc_feature
; NAME/KEY: misc_feature
; LOCATION: (21).(21)
; OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-184

Query Match 8.1%; Score 295.6; DB 19; Length 631;
Best Local Similarity 71.7%; Pred. No. 6; 8e-67;
Matches 407; Conservative 0; Mismatches 149; Indels 12; Gaps 1;
Qy 1708 ATGATGTTGTTGCTGAGGTTGCACTTGCTGTTGCTGAGCTTACCCAGCAATTCAGCAGAG 1767
Db 64 ATGACTGCACTGAGGTTGCACTTGCTGTTGCTGAGCTTACCCAGCAATTCAGCAGAG 123
Qy 1768 TCATTCAGTGTGCTGAGGTTGCACTTGCTGTTGCTGAGCTTACCCAGCAATTCAGCAGAG 1827
Db 124 AGCTTGAGGTATGGTGGCTGCTGAGCTTACCCAGCAATTCAGCAGAG 183

Qy 1828 CCACACCTTTTACAGATGCGATTAAGCAATTCCAGCAACTTCAGGGCTAAGTAAT 1887
Db 184 CCTCATCACATTCAAGATGCGAATACCTATTCAGAACCTCAGAACCTCAGAGTCAAGTAAC 243
Qy 1888 GCTATGGAGATTCTGGATGTGTATGATCACGGATAATTGGGAGTAGACTCGGT 1947
Db 244 GCTGTGACATTCCTGACAACTATACCAAGAAGAACGAGGAGA 303
Qy 3330 AGTGTGACAGTCAGCAGATGAAACCGTGCCTGGGAGCACTTGGAGCT 307
Db 1948 CTAGAAGGGGCTCTGCACTGGCTTATGGCAAAGTTGATGAAATGCCGATATGTTG 2007
Qy 304 TGGAGAGGGGACTCTTCACCTCTCTAGGGACCTTGATGACAGTCGTCITGTTG 363
Db 2068 ATTCATTAATCGTGTG-----ACATGATGATCCTCGACTATGCAATATG 2115
Qy 424 ACTCAAGGATGACATGACATGACATGACATGACATGACATGACATGACATGAC 483
Db 2116 TGGAAACCTGGTGGAGGSGGTTGCTTCTGATGTTCTGAGACATGCAATG 2175
Db 484 TGGAGACGTGTTGATGGAGCTGTTGATTCCTCCAGGTTAGACACACAGATAG 543
Qy 2176 TTAAACTGGGACACTATGATGATGATGATGATGATGATGATGATGATGATG 2235
Db 544 TTAGACTGGGAGACTATGATGATGATGATGATGAGTACTAGAGCTGGAT 603
Qy 2236 GTAGTCGGGTCCTTGTGCTG 2263
Db 604 GGCACTATGGTCACCTGACTGCTG 631

RESULT 10
US-10-424-599-129007
; Sequence 129007 Application US/10424599
; Publication No. US20040131072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Plant Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 3B-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 129007
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE: misc_feature
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87499C.1
US-10-424-599-129007

Query Match 7.2%; Score 262.4; DB 18; Length 1146;
Best Local Similarity 58.7%; Pred. No. 5; 8e-58;
Matches 630; Conservative 0; Mismatches 266; Indels 178; Gaps 3;
Qy 587 TCTGCTCGGCGGCAAATGGCCGACCGCTCTCCGACTCTCAATTCACTCGGATT 646
Db 236 TATCCGCACTGGCAATGGGGGGGGCTCTGGATTCGATTCCTGGGAGC 295
Qy 647 CCTCCCTCTCTCTGGCACCGACACCCACCTCTCGCTCTCGGACCAT 706
Db 296 CGCCCGCTCCTCT-----CCACCTCCACCTCTCGCCCTCTCGTCCCTCCC 346
Qy 707 CTATTGATGCGCCGAGCGCCAGTCCATCCCAATGATTCTACCAAGTATTAGGAG 766
Db 347 GCCTCGATCTCTGAGGCTCTGCTGATCCCTCTGACCTGATCCCTCGGG 406

```

QY 767 CTCAAACCATTTAACCGATGAACTCAGAAGAACATCGAGCTGGTTTCGAAC 826

Db 407 CCACACCCACTTCTCGCGACGGCATTCGAGGGCTCAAGGTCTCGAAGC 466

QY 827 CGCGCAATTGGTTTCAAGGAGACGGCTTAATCAGCGGAGACAGATCTCAAGTG 886

Db 467 CTCTCAGTAGCTTCAAGAAGACGTTTAATGGCGCCCAAATCTCGAAGC 526

QY 887 CTCTCGGAAACCTCTAGCTGATCTTCCAGAAGAGACTACATGAAAGTCCTTG 946

Db 527 AACGAGAGGCCTACCTCAATTCTCGAAGAAGAGTACATCAAAAGCCTTGCGACG 586

QY 947 ATCGAGAGACTCTGCTTAATCCCTGGCTTAAAGAGAGACTACATGAAAGTCCTTG 946

Db 587 ACCAGAGAGCCGCTACCTCAATTCTCGAAGAAGAGTACATCAAAAGCCTTGCGACG 586

QY 1007 AATAAAGTTCTCGTTTAATTCTGATGAAATGATAAAGGAAGACTTTTACTAGT 1066

Db 620 -----TCGAC 624

QY 1067 AACGTTCTGGGACTCTCTGTTGATGCGAGAAGGTTGAGATAGSTCTCTCGG 1126

Db 685 ATTGGGGGGGTTGCTTGGAGGAGGGTGCCTAACAGAGCTTACGGATGTTG 744

Db 625 AACGTTCTGGACGGCTGTTGTTGAGCTGGAGAGCTGGAGAGCTGGCTG 684

QY 1127 GTTGTGGGGCTCTGCTTAAGGAGAGGTGCTTAACTGAGATGTTGGTTA 1186

Db 745 GATATGGGACTCCATTGTTGAGCTGTCAGGAGTCTGACCTTCTCCACCGAT 804

QY 1187 GTTATGCGGCTCTGTTCTCGATGCTGAGGGATGCTATGCCATCTGATCCACCTGAT 1246

Db 805 TCTATGCGGCTCTGAGATGCTGAGGGCTTGAAGCTTGGAGA----- 855

QY 1307 GCTTGGGAACTTGTGAGGGCTTGTGACGAGACTCTCTGACCTTGTTA 1366

Db 856 -----GAAGGGGAAACCGAGCTAGTCGACGTTAACGACAAATA 897

QY 1427 GATGAGACTTGTGAGGAGATCACTCCSGTTAGCTCTGGAGCTACTGGCTTACCGCTT 1486

Db 898 GATGAGAGCGCTACAAAGAGATAACCCACGGTTGTTGAACTTGTGCTGCCTT 957

QY 1487 GGTGATGATACCTGGAAAGACTAAATGGTTAACGGGTGGGGAAATTGTTG 1546

Db 958 GATGAGAGACATCTGAGCCGGGAGGGAGCTTCTGTTGCTCACATTG 1017

QY 1547 TCGTTGGGGGGGTGGAGCATGCTCTGTTGGGGTTGACCCGTGAGAGTTATG 1606

Db 1018 GCGTTGGGGGGGGGAGAACAAATTGGGGGTTAACCGGTGAGACTCTG 1077

QY 1607 ATGAGGGGTTTATGAGCTGAGCTGACGAGCTGTTACGGTTAGATAC 1660

Db 1078 AATGAGGCACTTACACGAGCAGGGCTGAGACGGTTGACTTTGTAGCC 1131

RESULT 11  
US-10-600-070-135

; Sequence 135, Application US/10600070

; Publication No. US20040139500A1

; GENERAL INFORMATION:

; APPLICANT: Osteryoung, Katherine W.

; APPLICANT: Vitha, Stanislav

; APPLICANT: Koksharova, Olga A.

; TIME OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of

; FILE REFERENCE: MSU-08153

; CURRENT APPLICATION NUMBER: US/10/600,070

; CURRENT FILING DATE: 2003-06-20

; NUMBER OF SEQ ID NOS: 206

; SEQ ID NO 135

; LENGTH: 660

; ORGANISM: Medicago truncatula

; US-10-600-070-135

; Query Match 5.7%; Score 208.8; DB 19; Length 660;

; Best Local Similarity 60.3%; Pred. No. 5; 8e-44;

; Matches 429; Conservative 0; Mismatches 207; Indels 76; Gaps 2;

; Db 561 ACCTGGCCACAAACCTCTACAACTATCTGCTCGCCAGAACATGGGCCACCGTCTCT 620

; Db 17 ACCTAACCCCTCCTACATTCTCCCGGCTCCGGCACCCAGATAATGGGCCAGGGACTCAT 76

; QY 621 CTCGACACTCAATTCTACCTCCGATTCCTCCCTCTCTCGCTCGCCAGAACATGGGCCACCGCACCAC 680

; Db 77 TCCGATTTCAAATTCTCTGGCACACTCTCTCCACCACTCCACCTCGGAC 136

; Db 681 CGCGACTCTGTCCTCTGCCACCATCTATGATGTCGGCGAAAGCCAGGTCCACCTCC 740

; Db 137 AGTCACCTC-----ACTCTCTTACCCCTCTCCGATAGAACGCCCGTCTGACTCC 190

; Db 741 CATGATGTTCTACCGAGGTTTAAAGAGCTTCACCGATGTTCTAACCGATGTTAGAG 800

; Db 191 TCTCGACCTGACAAATCTCTCGCCGGCGAACAGCATTCTCGGTGAGTTGG 250

; QY 801 AGCATTCGAGCTGGGTTCGAAACCCCCGAAATCGGTTAGCGAGGACCTTAA 860

; Db 251 ASCITATGAGCAAAATCTCTGAGCCCTCTCAGTATGCTTCAGTAATGAGCTTGT 310

; Db 861 CAGCGGAGACAGATCTCAAGCTGCTGCGAAACTCTGCTAATCCTGGCTAGAG 920

; Db 311 TACGTGTCAGATCTCAAGCTGCTGCTGATGAACTGATCTGCTCTAGAG 370

; QY 921 AGAGTACATGAGGTCCTCTGATGATGAGAGCTAACGTCAACTGATGTTCTCTG 980

; Db 371 AGATGATATCACAGCCTGTCGCGAGAG----- 405

; QY 991 GGATAGGTAAATTGTTGGAATAATAAAGTTCTCGTTTAATTCTGAATGG 1040

; Db 406 -----AAGATGAGGAATCT 420

; QY 1041 ATAAGGAGGAACTTTATCTAGTGAAGGTCTGGGGCTCTCTGTTGTTGAGAG 1100

; Db 421 CCTTCCTCTGAAATCCCTTGCACAAAGTCTGAGCTCTGTTGTTGAGAG 480

; QY 1101 CTGGTGGAGCTGAGATGACTCTCTGGGGTTGGAGGCTCTGCTTAAGGAAGGGTTGCTA 1160

; Db 481 CTGGAGAGGGGGTTGGCTCTGGGGTTACTGAGAGGAGGGTTACCGA 540

; QY 1161 ACTGGTTTACCGAGAGCTGGTTAGTATGGGGCTGGCTCTGATGTTGCTGGGG 1220

; Db 541 AGATGTTTACCGAGAGCTGGTTGGCTCTGGGGCTCTGATGTTGAGGTTCTAGGG 600

; QY 1221 ATGCTATGGATGATGACCTGATTACTGGTTAGTGGTTGAGTTG 1272

; Db 601 ATCTATGGCTTGTCCCCCGAGATTCATTGCTGTTGAGATGCTGG 652

RESULT 12  
US-10-600-070-126

; Sequence 126, Application US/10600070

; Publication No. US20040139500A1

; GENERAL INFORMATION:

; APPLICANT: Osteryoung, Katherine W.

; APPLICANT: Vitha, Stanislav

; APPLICANT: Koksharova, Olga A.

; TIME OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of

; FILE REFERENCE: MSU-08153

; TITLE OF INVENTION: Use  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; CURRENT FILING DATE: 2003-06-20  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 126  
; LENGTH: 2283  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; US-10-600-070-126  
; Query Match 5.3%; Score 194; DB 19; Length 2283;  
; Best Local Similarity 55.6%; Pred. No. 1e-39; Mismatches 300; Indels 110; Gaps 3;  
; Matches 514; Conservative 0;  
; 1368 TGAGTCCTGTRGGAGGAGGAGCAAGTAGCCTTGACCGGATTAGTGCAATTG 1427  
; 716 TCAAGCTTGCAGGAGAATGGCAAGCAATTCGCAACCTGTATCTCTTCAGATG 775  
; 1428 ATCAGACTTGGAGAGAGATCACTCCGCGTTATCTTGAGAGCTACTGGCTTAACCGCTG 1487  
; 776 ATGAAACTCTCGAGGAGATTACACCTCGCTGTGTTATGGAGCTCTCCCTTCTATTG 835  
; 1488 GTGATGTTACGAGCTGGAAAAGCTAAATGGTTAASCGGTGCGGSAATTGGGT 1547  
; 836 ACAGAGCATCATAGAGGCAAGGCGCAAGAGGGCTCAAGGTGCGAGA 895  
; 1548 CTGTTGGAGGAGGTGGAGCATCAGCTCTGTTGGGGTTGACCGGTTGAGAGTTATGA 1607  
; 896 GCCTTGCGAGAGGATATGCTTACCGTGTGAGGAGGTTTCTGTCAGCGCTCATGA 955  
; 1608 ATAGGGCGTTTACGAAATGACGCTGCTGAGGAGGTTACAGTTAGATAACCTTTT 1667  
; 956 ACCBAGCTTTCAGGATGACAATGAAAGATG----- 993  
; 1668 AATTTCTTGTAGCATGATAAACCTTGTGTTCTCATTTAATGATGTTGTTGTTGAGT 1727  
; 994 ----- 993  
; Db 1728 TGATCTTTGCTACCCANGCAATTCCAGAGTCATTGAGTTACGAGT 1787  
; 994 -GATTCCTTTCAAACACCGGAATAGCTTCCATGTTGAAATTACAGT 1052  
; Db 1788 TGCAGTCCTGTTGGCTCAAGCTTATGGTAGAGGAGCCACCTTTAGGGATG 1847  
; 1053 AGGACTTCACATCTGCTCAAGCAATTAGTTAAAGGCCCAATTCTGATGCG 1112  
; 1848 TGTAAASGATTCAGCAACTTGTGGAGCTAGCTTGGTAATGGTATGGCTATGGCT 1907  
; 1113 GGATGATCTTTGAACTCTCGAGGTTCAACATAGTTCTCATATGCT----- 1164  
; Db 1908 GTTGTATGATCACCGGAATAATGGAGATGCTTGGCTCTGAGAAMGGGGCTCTGGC 1967  
; 1165 -----TATGATATGATGAGATGACCTCTGATGAGAAGGGCTCTGCTC 1208  
; Db 1968 ACTCTTCTAGGCAAGTGTGATGAGCTGGCTATGGTTGGGGTTAGCAGTGGAGTC 2027  
; 1209 ATTGCTACTCGGGAGATGTAGCAGTGCAGAATGTCGCTGATGATGATGCTC 1268  
; Qy 2028 ACATATAGGAATCCAGTATTGGAGTTGTGTTGGAAACCTGGTGGCAGGGTTGCTT 2084  
; Db 1269 ACCATACAGACCCAAATTCTAGGTTATGTGACCAACTCTAGCATCTGAGA 1328  
; Qy 2085 CATTGATCATCTCCCTGCACTAGTGCACATTGGTGGCAGGGTTGCTT 2144  
; Db 1329 GATGATCTCTTCAGGCTGGCAACTTGTGAGCTGCTTATCTTGGGTT 1388  
; Qy 2145 TCTAGGTTGAGGTTACTTGGAAAGAGTGGAGGAGTTGTCAGGGTCTCCCTTACGTC 2204  
; Db 1389 TCTAGGAGCAGGAGATCTCGGGCATCGAGTCAGCTGAGGATGATGATGCC 1448  
; Qy 2205 TAGGTTTGTAGTTACTTGGAAAGAGTGGAGGAGTTGTCAGGGTCTCCCTTACGTC 2264

RESULT 13  
; US-10-021-323-3562  
; Sequence 3562, Application US/10021323  
; Publication No. US2004012334011  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; Title Reference: 30-1(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIORITY APPLICATION NUMBER: US 6/0/255, 619  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO: 3562  
; LENGTH: 607  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1). (607)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: LIB3825-027-Q6-K6-H1  
; US-10-021-323-3562

Query Match 4.9%; Score 178.6; DB 19; Length 607;  
; Best Local Similarity 69.3%; Pred. No. 5.2e-36; Mismatches 100; Indels 18; Gaps 1;  
; Matches 266; Conservative 0;  
; 1986 TGATGATGCCCCPATGTTGTTGGCTTAGACGTGAGGATCAAAATGAAATCCAGC 2045  
; 5 TGNAGCTGGCGTGTGGTTGGCTTAGACGTGATAGTAGCTCCCTGAGCT 64  
; Db 2046 TATGTGGGTTGTTGAGGATCAATCTGATGAGATCTCCCTGGACT 2105  
; 65 TATGTGAGATTGCTGTTGAGAATCTAAAGCATGAGCTGAGAGATCTCCCGGT 124  
; Qy 2106 ATGCAAAATGTTGAGAACCTGGTGGAGGGGTGTCTCCCTAGGTCTAGACCA 2165  
; 125 TTGCAACTCTGGAGGCTATGGAGTTCTGAGTTAGACACCA 184  
; Db 2166 AGTAAATAATTAAACTCTGGGACTACTATGATGATCTATGGTTGGAGTACTTGA 2225  
; Qy 2185 AGATATACTTAACTGTTGGAGATTATGATGATCTACTGCTGAGTATTGA 244  
; Db 2226 AAGAGTGGGGAGTTGAGGTTCTCTTACTGCGCTGCTGAACTATGCAAGGATG 2285  
; 245 AAGCTTGTGAGGGAGAGGGTGTACCCCTTGTCTGAGCTGAGCTATGAGGATAG 304  
; Qy 2286 AGCGAG-----CATGTAAGCTAGTGTGTTGAGGCACTGGAGA 2327  
; Db 305 TGTGAGGCTACTCTGAGTTCTGTGATGATGAGCTGAGGATGAGGATGAGA 364  
; Qy 2328 AGTTTCTCCCGCGTATCAGA 2351  
; Db 365 GGTGTTCTCTGGCTCAGA 388

RESULT 14  
; US-09-732-627A-2154  
; Sequence 2154, Application US/09732627A  
; Publication No. US20040123339A1

; GENERAL INFORMATION:  
; APPLICANT: Fischer, Karen L.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(51770)B  
; CURRENT FILING DATE: 2000-12-08  
; SEQ ID NO 2154  
; LENGTH: 439

; ORGANISM: *Gospypium hirsutum*  
; FEATURE: NAME/KEY: unsure  
; LOCATION: (1)..(439)  
; OTHER INFORMATION: unsure at all n locations

US-09-732-62A-2154

Query Match 4.8%; Score 175; DB 11; Length 439;  
Best Local Similarity 65.0%; Pred. No. 3.7e-35; Matches 278; Conservative 0; Mismatches 141; Indels 9; Gaps 1;

Qy 499 GTCGCGATGGTGTCTCCGCATTCGAATTATGGCGATTACACGGCAGCAAAAGCTC 558  
Db 12 GTAGCATCGGGGTGCACTCCACACTGCTGCTCTTATGACCCCTGTAAACCTTC 71  
Qy 559 CGAGGTAGGCCAACACCTCTACAACATGTGCGCGAGCAATGGCCGACCGCTCT 618  
Db 72 ANATCCACACCTCTACAACGCTGCTGCGCCGAAATGGCCGAGGCGCTCT 131  
Qy 619 CTCTCGGACTCTAATTCTACCTCG-----ATTCTCTCTCTCTCTCGCCACC 669  
Db 132 CTCTCTGATCTCAATTCTCCCTGCCCGATTAATCGTCCTCTCTCTCTCTCTCA 191  
Qy 670 GCACCAACACCCSACACTCTCGCTCTCGCCACATCTATGCGCCAGCACGCCAC 729  
Db 192 ACGCCACACTCTTCCTCCCTTAACCTCCCTCTCGCCCTCTCCCGAACGCCAC 251  
Qy 730 GTCGGCATCCCATGATTCCTCCAGGTTAGGGAGCTCAAAACACATTCCTAACCG 789  
Db 252 GTCCTCATTCCTCTCGCTCTCGCCAGCTGAGCTGAGACTCATTCCTTCTGCT 311  
Qy 790 GGATCAGAGAGCATGAGCTGAGGTTCGAACCCCGCATTCGGTTAGCGAC 849  
Db 312 GGATTTAGAGAGACGCTTATGAGCAGGGTTGCAACCGCTCATATGGTTAGCTCA 371  
Qy 850 GAGCTTAAATGCCGAGACGATCTCAAGCTGCTTGCAAAACTCTGTATACCT 909  
Db 372 GACCATTAATTAGCCGAAGACGATTCCTACTGCTGCTGAAACCTATCTAACCT 431  
Qy 910 CGCTCTAG 917  
Db 432 GCGCTCTAG 439

RESULT 15  
US-10-600-070-175  
; Sequence 175, Application US-10-600070  
; Publication No. US2004013950A1  
; GENERAL INFORMATION:  
; APPLICANT: Oberyoung, Katherine W.  
; APPLICANT: Vaithe, Stanislav  
; APPLICANT: Kokharcova, Olga A.  
; APPLICANT: Gao, Hongo  
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
; TITLE OF INVENTION: Use  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; CURRENT FILING DATE: 2003-06-20  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 175

Length: 545  
TYPE: DNA  
ORGANISM: *Gospypium arboreum*  
FEATURE: NAME/KEY: misc feature  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE: NAME/KEY: misc feature  
LOCATION: (528)..(528)  
OTHER INFORMATION: OTHER INFORMATION: n is a, c, g, or t  
FEATURE: NAME/KEY: misc feature  
LOCATION: (540)..(540)  
OTHER INFORMATION: n is a, c, g, or t  
US-10-600-070-175

Query Match 4.4%; Score 161; DB 19; Length 545;  
Best Local Similarity 69.6%; Pred. No. 2.1e-31; Matches 218; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 3224 ATAGGTTTGGATGGCGGAATGCTGAAGTTGGACTGAGAGCTGAAACTGGCGA 3283  
Db 82 AGAGCTTCGGATGGTCAATGTTGAACTGAGAGCTGAGAGCTGAGCTC 141  
Qy 3284 GCTTGGGGTGTATGATACTGTTGAAACTATCTGTGACAGTGTGACAGTC 3343  
Db 142 GCTTGGGGTGTATGATACTGTTGAACTGAGAGCTGAGAGCTGAGCTC 201  
Qy 3344 AGAGATGCAACCCGGCTGCTGTTGTTGGAAGCAACTCTGGAGACTCTGCTGTCATCTGCA 3403  
Db 202 ACTGATGGCCACGGAGCTGAGCTACTCTGGAGAGCTGAGCTGAGCTGAGCA 261  
Qy 3404 TTGGITCTCCAGAAACAGTGTACTCTGTCAGAACTACACAGAGATAGCTGAGCT 3463  
Db 262 TGTGTCATCTCCGAGAAACATGCTCTGTAACCTCTGAGATGAGAT 321  
Qy 3464 TTTCGGGCAAGTCAGGGGAAATCTGTGAGGGCTCTGTCATCATATAC 3523  
Db 322 GCTCTGTCACCTCAGGTGGAAATCTGTGCTACAAATCTRACTAG 381  
Qy 3524 TCATATGGCAT 3536  
Db 382 ATGTTAAAGCAT 394

Search completed: June 10, 2005, 01:37:37  
Job time : 2087 sec

THIS PAGE BLANK (USPTO)







Query Match Score 50.2; DB 4; Length 57280;  
 Best Local Similarity 53.9%; Pred. No. 0.031; Mismatches 88; Indels 0; Gaps 0;  
 Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

US-09-949-016-13543 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 13543 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human

**RESULT 7**  
 US-09-949-016-13542/C  
 Sequence 13542, Application US/09949016  
 ; GENERAL INFORMATION:  
 ; Parent No. 6812339  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CI001307  
 ; CURRENT APPLICATION NUMBER: US/09/949, 016  
 ; PRIOR APPLICATION NUMBER: 60/241, 755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237, 768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231, 498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 13542  
 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; US-09-949-016-13542

Query Match Score 50.2; DB 4; Length 57280;  
 Best Local Similarity 53.9%; Pred. No. 0.031; Mismatches 88; Indels 0; Gaps 0;  
 Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

US-09-949-016-13544/C ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 13544 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human

**RESULT 8**  
 US-09-949-016-13543/C  
 Sequence 13543, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CI001307  
 ; CURRENT APPLICATION NUMBER: US/09/949, 016  
 ; PRIOR APPLICATION NUMBER: 60/241, 755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237, 768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231, 498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 13544  
 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; US-09-949-016-13544

Query Match Score 50.2; DB 4; Length 57280;  
 Best Local Similarity 53.9%; Pred. No. 0.031; Mismatches 88; Indels 0; Gaps 0;  
 Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

US-09-949-016-13545 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 13545 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human

**RESULT 9**  
 US-09-949-016-13544/C  
 Sequence 13544, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CI001307  
 ; CURRENT APPLICATION NUMBER: US/09/949, 016  
 ; PRIOR APPLICATION NUMBER: 60/241, 755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237, 768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231, 498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 13544  
 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; US-09-949-016-13544

Query Match Score 50.2; DB 4; Length 57280;  
 Best Local Similarity 53.9%; Pred. No. 0.031; Mismatches 88; Indels 0; Gaps 0;  
 Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

US-09-949-016-13546 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 13546 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human

**RESULT 10**  
 US-09-949-016-13545/C  
 Sequence 13545, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CI001307  
 ; CURRENT APPLICATION NUMBER: US/09/949, 016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241, 755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237, 768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231, 498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012

RESULT 10  
 US-09-949-016-13545/C  
 ; Sequence 13545, Application US/09949016  
 ; Patient No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIORITY NUMBER: 60/241,755  
 ; PRIORITY FILING DATE: 2000-10-20  
 ; PRIORITY NUMBER: 60/237,768  
 ; PRIORITY FILING DATE: 2000-10-03  
 ; PRIORITY APPLICATION NUMBER: 60/231,498  
 ; PRIORITY FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 13545  
 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-13545

Query Match 1.4%; Score 50.2; DB 4; Length 57280;  
 Best Local Similarity 53.9%; Pred. No. 0.031; Mismatches 88; Indels 0; Gaps 0;  
 Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```

Qy      565 AGCCACAAACCTTACACTACTATGCTCGGCCAGAATGGCCGACCGTCCTCTC 624
Db     13101 AGCACCATACCAACACTGCCTCCACCATCACATTGAGTC 13042
Qy      625 GACTTCATAATTCACTCCATTCTCTCTCTGCCCCACCGCC 684
Db     13041 ACCATCAATACCTATCATCGTCCACCAACCGCTTACCAT 12982
Qy      625 ACTCTCGTCTCTGCCACCATCATGTTGCGAACGGCTTACCA 744
Db     12981 ACACACCACACTACTATCATGACTTCCACCATACAC 12922
Qy      745 GATTCTTACCA 755
Db     12921 CATCATCAGGA 12911

```

RESULT 11  
 US-09-949-016-14634/C  
 ; Sequence 14634, Application US/09949016  
 ; Patient No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIORITY NUMBER: 60/241,755  
 ; PRIORITY FILING DATE: 2000-10-20  
 ; PRIORITY NUMBER: 60/237,768  
 ; PRIORITY FILING DATE: 2000-10-03  
 ; PRIORITY APPLICATION NUMBER: 60/231,498  
 ; PRIORITY FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 14634  
 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-14634

Query Match 1.4%; Score 50.2; DB 4; Length 57280;  
 Best Local Similarity 53.9%; Pred. No. 0.031; Mismatches 88; Indels 0; Gaps 0;  
 Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```

Qy      565 AGCCACAAACCTTACACTACTATGCTCGGCCAGAATGGCCGACCGTCCTCTC 624
Db     13101 AGCACCATACCAACACTGCCTCCACCATCACATTGAGTC 13042
Qy      625 GACTTCATAATTCACTCCATTCTCTCTCTGCCCCACCGCC 684
Db     13041 ACCATCAATACCTATCATCGTCCACCAACCGCTTACCA 12982
Qy      625 ACTCTCGTCTCTGCCACCATCATGTTGCGAACGGCTTACCA 744
Db     12981 ACACACCACACTACTATCATGACTTCCACCATACAC 12922
Qy      745 GATTCTTACCA 755
Db     12921 CATCATCAGGA 12911

```

RESULT 12  
 US-09-949-016-14634/C  
 ; Sequence 14634, Application US/09949016  
 ; Patient No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIORITY NUMBER: 60/241,755  
 ; PRIORITY FILING DATE: 2000-10-20  
 ; PRIORITY NUMBER: 60/237,768  
 ; PRIORITY FILING DATE: 2000-10-03  
 ; PRIORITY APPLICATION NUMBER: 60/231,498  
 ; PRIORITY FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 14634  
 ; LENGTH: 57280  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-14634

Query Match 1.4%; Score 50.2; DB 4; Length 57280;  
 Best Local Similarity 53.9%; Pred. No. 0.031; Mismatches 88; Indels 0; Gaps 0;  
 Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```

Qy      565 AGCCACAAACCTTACACTACTATGCTCGGCCAGAATGGCCGACCGTCCTCTC 624
Db     13101 AGCACCATACCAACACTGCCTCCACCATCACATTGAGTC 13042
Qy      625 GACTTCATAATTCACTCCATTCTCTCTCTGCCCCACCGCC 684
Db     13041 ACCATCAATACCTATCATCGTCCACCAACCGCTTACCA 12982
Qy      625 ACTCTCGTCTCTGCCACCATCATGTTGCGAACGGCTTACCA 744
Db     12981 ACACACCACACTACTATCATGACTTCCACCATACAC 12922
Qy      745 GATTCTTACCA 755
Db     12921 CATCATCAGGA 12911

```

RESULT 13  
 US-09-949-016-14635/C  
 ; Sequence 14635, Application US/09949016  
 ; Patient No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: Fast-SEQ for Windows Version 4.0

SEQ ID NO: 14635

LENGTH: 57280

TYPE: DNA

ORGANISM: Human

US-09-949-016-14635

Query Match

Best Local Similarity 53.9%; Score 50.2; DB 4; Length 57280; Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 565 ASCCACACACTCTTCAACTATCTGCGCCAGAAATGGGCCACCGTCCTC 624  
Db 13041 ACCATCAATTCCTATCAGTCGCCCCACCCAGGTCTTCACCATCACCC 13042

Qy 625 GACTTCATTTCACCTCGATCTCTCTCTGCGACCGACCCGG 684  
Db 13041 ACCATCAATTCCTATCAGTCGCCCCACCCAGGTCTTCACCATCACCC 12982

Qy 685 ACTCTCGTCTCTGCCACCATCTATGATCGRCCCCAACGGCCACGTC 744  
Db 12981 ACACACCACTCACTATCATGACTTCCACCATCACTCCACCAAC 12922

Qy 745 GATTCTTACCA 755  
Db 12921 CATTACAGA 12911

RESULT 15

US-09-949-016-14637/c

Sequence 14637 Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: Fast-SEQ for Windows Version 4.0

SEQ ID NO: 14637

LENGTH: 57280

TYPE: DNA

ORGANISM: Human

US-09-949-016-14637

Query Match

Best Local Similarity 53.9%; Score 50.2; DB 4; Length 57280; Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 565 ASCCACACACTCTTCAACTATCTGCGCCAGAAATGGGCCACCGTCCTC 624  
Db 13041 ACCATCAATTCCTATCAGTCGCCCCACCCAGGTCTTCACCATCACCC 13042

Qy 625 GACTTCATTTCACCTCGATCTCTCTCTGCGACCGACCCGG 684  
Db 13041 ACCATCAATTCCTATCAGTCGCCCCACCCAGGTCTTCACCATCACCC 12982

Qy 685 ACTCTCGTCTCTGCCACCATCTATGATCGRCCCCAACGGCCACGTC 744  
Db 12981 ACACACCACTCACTATCATGACTTCCACCATCACTCCACCAAC 12922

Qy 745 GATTCTTACCA 755  
Db 12921 CATTACAGA 12911

Search completed: June 10, 2005, 01:02:51

Job time : 583 secs

Query Match 1.4%; Score 50.2; DB 4; Length 57280;  
Best Local Similarity 53.9%; Pred. No. 0.031; Mismatches 88; Indels 0; Gaps 0;  
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 565 ASCCACACACTCTTCAACTATCTGCGCCAGAAATGGGCCACCGTCCTC 624  
Db 13041 ACCATCAATTCCTATCAGTCGCCCCACCCAGGTCTTCACCATCACCC 13042

Qy 625 GACTTCATTTCACCTCGATCTCTCTCTGCGACCGACCCGG 684  
Db 13041 ACCATCAATTCCTATCAGTCGCCCCACCCAGGTCTTCACCATCACCC 12982

Qy 685 ACTCTCGTCTCTGCCACCATCTATGATCGRCCCCAACGGCCACGTC 744  
Db 12981 ACACACCACTCACTATCATGACTTCCACCATCACTCCACCAAC 12922

Qy 745 GATTCTTACCA 755  
Db 12921 CATTACAGA 12911

LENGTH: 57280

TYPE: DNA

ORGANISM: Human

US-09-949-016-14636

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 18:54:19 ; Search time 1797 Seconds  
(without alignments)

12079.962 Million cell updates/sec

Title: US-10-600-070-3

Perfect score: 3667  
Sequence: 1 tggtctgcattaaggagaa.....ctataacatcaaagggtacaa 3667

Scoring table: IDENTITY\_NUC

Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseqm1980s:\*

2: geneseqm1990s:\*

3: geneseqm2000s:\*

4: geneseqm2001bs:\*

5: geneseqm2002bs:\*

6: geneseqm2002as:\*

7: geneseqm2002bs:\*

8: geneseqm2003as:\*

9: geneseqm2003bs:\*

10: geneseqm2003cs:\*

11: geneseqm2003ds:\*

12: geneseqm2004as:\*

13: geneseqm2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description	RESULT 1	ALIGNMENTS
1	3667	100.0	3667 12 ADJ38130	ADJ38130 Arabidopsis	ID ADJ38130 standard; DNA; 3667 BP.	
2	3663.8	99.9	3667 12 ADJ38136	ADJ38136 Arabidopsis	ID ADJ38130 Arabidopsis	
3	1931	52.7	2637 12 ADJ38210	ADJ38210 Plastid d	ID ADJ38130 Arabidopsis	
4	1724	47.0	2405 12 ADJ38129	ADJ38129 Arabidopsis	ID ADJ38130 Arabidopsis	
5	1720.8	46.9	2406 12 ADJ38135	ADJ38135 Arabidopsis	ID ADJ38130 Arabidopsis	
6	1717.6	46.8	2406 12 ADJ38208	ADJ38208 Plastid d	ID ADJ38130 Arabidopsis	
7	402.2	11.0	561 12 ADJ38212	ADJ38212 Plastid d	ID ADJ38130 Arabidopsis	
8	295.6	8.1	631 12 ADJ38264	ADJ38264 Plastid d	ID ADJ38130 Arabidopsis	
9	208.8	5.7	650 12 ADJ38215	ADJ38215 Plastid d	ID ADJ38130 Arabidopsis	
10	194	5.3	2283 12 ADJ38217	ADJ38217 Plastid d	ID ADJ38130 Arabidopsis	
11	178.6	4.9	607 13 ADJ48781	ADJ48781 Cotton pr	ID ADJ38130 Arabidopsis	
12	175	4.8	439 12 ADP3143	ADP3143 Cotton ex	ID ADJ38130 Arabidopsis	
13	161	4.4	545 12 ADJ38255	ADJ38255 Plastid d	ID ADJ38130 Arabidopsis	
c	159.4	4.3	552 13 ACM48855	ACM48855 Cotton pr	ID ADJ38130 Arabidopsis	
15	158	4.3	608 12 ADJ38217	ADJ38217 Plastid d	ID ADJ38130 Arabidopsis	
16	145.6	4.0	527 12 ADJ38214	ADJ38214 Plastid d	ID ADJ38130 Arabidopsis	
17	136.2	3.7	647 12 ADJ38265	ADJ38265 Plastid d	ID ADJ38130 Arabidopsis	
18	134.6	3.7	652 12 ADJ38266	ADJ38266 Plastid d	ID ADJ38130 Arabidopsis	
c	19	3.6	563 12 ADJ38258	ADJ38258 Plastid d	ID ADJ38130 Arabidopsis	
c	20	3.6	446 12 ADJ38231	ADJ38231 Plastid d	ID ADJ38130 Arabidopsis	

CC compositions of the present invention are useful for further characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is that of a DNA sequence which is related to the invention.

**XX** Sequence 3667 BP; 983 A; 670 C; 842 G; 1172 T; 0 U; 0 Other;

Query Match 100.0%; Score 3667; DB 12; Length 3667;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 3667; Conservatve 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTCTCGATTAAGGAGATACTATATTAGCAATTGGCTGTTGATTCACAGATTG 60  
Db 1 TGTCTCGATTAAGGAGATACTATATTAGCAATTGGCTGTTGATTCACAGATTG 60

Qy 61 CTGGCTATAGGATTCATGGCTGTGTTACATTAGTACGTCATAATAGTTG 120  
Db 61 CTGGCTATAGGATTCATGGCTGTGTTACATTAGTACGTCATAATAGTTG 120

Qy 121 AATTATCACATTCACTTGAGTTGAGTTAAGAAGAAGGGAAATGATGGGTTTGG 180  
Db 121 AATTATCACATTCACTTGAGTTGAGTTAAGAAGAAGGGAAATGATGGGTTTGG 180

Qy 181 TTAACTTAAGTACTCAAGATACTGTTACTGCTATATGTPA 240  
Db 181 TTAACTTAAGTACTCAAGATACTGCTATATGTPA 240

Qy 241 ATAGAGGAACTCCAACCGTTTAGGGAAAGATTTAGTACGTTACATCAT 300  
Db 241 ATAGAGGAACTCCAACCGTTTAGGGAAAGATTTAGTACGTTACATCAT 300

Qy 301 AAATGCCGCGCAACTGTACATCATAGATGTTTATTTTCAGTAGGGTGT 360  
Db 301 AAATGCCGCGCAACTGTACATCATAGATGTTTATTTTCAGTAGGGTGT 360

Qy 361 GTTTTGATTACTTACTCTAAATCAAATTCTAAACCTAGAGGACCAA 420  
Db 361 GTTTTGATTACTTACTCTAAATCAAATTCTAAACCTAGAGGACCAA 420

Qy 421 CAGTCTCTAACATGAAACAGAACAGTTTGTAGTAGCCTAAAGCACTCC 480  
Db 421 CAGTCTCTAACATGAAACAGAACAGTTTGTAGTAGCCTAAAGCACTCC 480

Qy 481 ATGGAAGCTCTGACTGTCAGTCGGATTGTTCTCCATTGCCATTCA 540  
Db 481 ATGGAAGCTCTGACTGTCAGTCGGATTGTTCTCCATTGCCATTCA 540

Qy 541 CGCGCGACACAGCTCGACCTAGCCACAACCTCTACACTATCTGCGCCGC 600  
Db 541 CGCGCGACACAGCTCGACCTAGCCACAACCTCTACACTATCTGCGCCGC 600

Qy 601 AAATGGCGACCTCTCTCTGACTCTAACCTCCATTCTCTCTCTCC 660  
Db 601 AAATGGCGACCTCTCTCTGACTCTAACCTCCATTCTCTCTCTCC 660

Qy 661 TTGCCACGCCACACGCCACCGCGACCTCTGCTCTCTGCCACACATCTTCTCC 720  
Db 661 TTGCCACGCCACACGCCACCGCGACCTCTGCTCTCTGCCACACATCTTCTCC 720

Qy 721 GAAGCCACGTCCTCCATGCCATGTATTCTACAGGTTAGAGCTAACACATT 780  
Db 721 GAAGCCACGTCCTCCATGCCATGTATTCTACAGGTTAGAGCTAACACATT 780

Qy 781 TTACCGTGGATCAGAGAGGATCTGAGCTAGGGTTGAAACCCGGCATCGT 840  
Db 781 TTACCGTGGATCAGAGAGGATCTGAGCTAGGGTTGAAACCCGGCATCGT 840

Qy 841 TTGCGGAGACCTTAACTCACCGGAGACAGATTCTCAACTGCTGCGAACCTG 900  
Db 841 TTGCGGAGACCTTAACTCACCGGAGACAGATTCTCAACTGCTGCGAACCTG 900

Qy 961 GTCATCACTGAGTCCCTGGATAMGTTGATTGGAATAATAGTTCTG 1020  
Db 961 GTCATCACTGAGTCCCTGGATAMGTTGATTGGAATAATAGTTCTG 1020

Qy 1021 GTTAAATTCTGATGATAAGGAAAGACTTTATCTAGTGAGGTTCTGGG 1080  
Db 1021 GTTAAATTCTGATGATAAGGAAAGACTTTATCTAGTGAGGTTCTGGG 1080

Qy 1080 GCTTAAGGAGGTGCTAACGTTAACGTTAGCTGTTAAGAGTGTCTGC 1120  
Db 1080 GCTTAAGGAGGTGCTAACGTTAACGTTAGCTGTTAAGAGTGTCTGC 1120

Qy 1141 GCTTAAGGAGGTGCTAACGTTAACGTTAGCTGTTAAGAGTGTCTGC 1140  
Db 1141 GCTTAAGGAGGTGCTAACGTTAACGTTAGCTGTTAAGAGTGTCTGC 1140

Qy 1201 GTTCTCGATGTCGAGGATCTGCTAACGTTAACGTTAGCTGTTAAGAGTGTCTGC 1260  
Db 1201 GTTCTCGATGTCGAGGATCTGCTAACGTTAACGTTAGCTGTTAAGAGTGTCTGC 1260

Qy 1260 GTTCTCGATGTCGAGGATCTGCTAACGTTAACGTTAGCTGTTAAGAGTGTCTGC 1260

Qy 1261 TGAGTTCTGAGGAACTTGAGCTTACAGCTTACAGGTTAGCTGTTAAGAGTGTCTGC 1320  
Db 1261 TGAGTTCTGAGGAACTTGAGCTTACAGCTTACAGGTTAGCTGTTAAGAGTGTCTGC 1320

Qy 1321 AGCAGCGTTGGCTTATAGAACGTTCTGTTGATTTAGCTGTTAAGAGTGTCTGC 1380  
Db 1321 AGCAGCGTTGGCTTATAGAACGTTCTGTTGATTTAGCTGTTAAGAGTGTCTGC 1380

Qy 1381 GGAGGAAGGAGCAGTACCTCCATTGCCATTAGCTGAAATGAGACTTGG 1440  
Db 1381 GGAGGAAGGAGCAGTACCTCCATTGCCATTAGCTGAAATGAGACTTGG 1440

Qy 1440 AGGATCACTCGGGTTATGCTTGAGGACTCTGGCTTACCGCTTGGATGAGTC 1500  
Db 1441 AGGATCACTCGGGTTATGCTTGAGGACTCTGGCTTACCGCTTGGATGAGTC 1500

Qy 1501 TCGAAAGACTAAATGTTAAGCGSCTGCGGAATTATTGTTGCTCTGGAGG 1560  
Db 1501 TCGAAAGACTAAATGTTAAGCGSCTGCGGAATTATTGTTGCTCTGGAGG 1560

Qy 1560 TCGAAAGACTAAATGTTAAGCGSCTGCGGAATTATTGTTGCTCTGGAGG 1560

Qy 1561 TGGAGCTCAGCTCTGTTGGGGTTGACCGCTGCGGAATTATTGTTGCTCTGGAGG 1620  
Db 1561 TGGAGCTCAGCTCTGTTGGGGTTGACCGCTGCGGAATTATTGTTGCTCTGGAGG 1620

Qy 1620 TGGAGCTCAGCTCTGTTGGGGTTGACCGCTGCGGAATTATTGTTGCTCTGGAGG 1620

Qy 1621 ACCAATGAGCTGCTGAGCAGGTTACGTTAGATACCTTTTAATTCCTTGA 1680  
Db 1621 ACCAATGAGCTGCTGAGCAGGTTACGTTAGATACCTTTTAATTCCTTGA 1680

Qy 1681 TGATATACTTGTGTTCTCATTTATGATGTTGTTGTTGTTGTTGTTG 1740  
Db 1681 TGATATACTTGTGTTCTCATTTATGATGTTGTTGTTGTTGTTGTTG 1740

Qy 1741 GCTACCCAGCAATTCAGCAGACTCTGTTGAACTCTGCTGCTGCTGCT 1800  
Db 1741 GCTACCCAGCAATTCAGCAGACTCTGTTGAACTCTGCTGCTGCTGCT 1800

Qy 1800 GTCACCCAGCAATTCAGCAGACTCTGTTGAACTCTGCTGCTGCTGCT 1800

Qy 1801 GTCACCCAGCAATTCAGCAGACTCTGTTGAACTCTGCTGCTGCTGCT 1860  
Db 1801 GTCACCCAGCAATTCAGCAGACTCTGTTGAACTCTGCTGCTGCTGCT 1860

Qy 1860 GTCACCCAGCAATTCAGCAGACTCTGTTGAACTCTGCTGCTGCTGCT 1860

Qy 1861 CAGCAACTTCTGAGCTTATGCTGAGGACCTTCTGAGCTGCTGCTGCT 1920  
Db 1861 CAGCAACTTCTGAGCTTATGCTGAGGACCTTCTGAGCTGCTGCTGCT 1920

Qy 1920 CAGCAACTTCTGAGCTTATGCTGAGGACCTTCTGAGCTGCTGCTGCT 1920

Qy 1921 CGGAAATTGGAGATAGACTCTGCTGAGGAGCCACCTTCTGAGCTGCTGCT 1980  
Db 1921 CGGAAATTGGAGATAGACTCTGCTGAGGAGCCACCTTCTGAGCTGCTGCT 1980

Qy 1980 CGGAAATTGGAGATAGACTCTGCTGAGGAGCCACCTTCTGAGCTGCTGCT 1980

Qy 1981 AAATGATGATGAGCTGCTGAGGTTAGACAGTGTAGGATTACATAGGAAT 2040



DR WPI: 2004-002486/08.  
 DR P-FSDB; ADJ38203.

XX New isolated Ftn2, ARCs and/or Fzo-like nucleic acid sequences useful  
 PT for further characterizing plastid division in plant cells, and in  
 PT varying agronomic and horticultural characteristics of economically  
 PT important plants.

XX Claim 1; SEQ ID NO 10; 287pp; English.

CC This invention relates to novel prokaryotic type or plastid division and  
 CC related genes and proteins. In particular, the invention relates to novel  
 CC compositions of the present invention are useful for further  
 CC characterising plastid division in plant cells, in order to vary  
 CC agronomic and horticultural characteristics of economically important  
 CC plants, such as crop, ornamental and woody plants. They can also be used  
 CC as herbicide targets. The present sequence is that of a DNA sequence  
 XX which is related to the invention.

SQ Sequence 3667 BP; 982 A; 669 C; 843 G; 1173 T; 0 U; 0 Other;

Query Match 99.9%; Score 3663.8; DB 12; Length 3667;  
 Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;  
 Matches 3665; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGTCTGCAATTAGAGAATACCAATTATAGCAATTGGCTTCAACAGATTG 60  
 Db 1 TGTCCTGCAATTAGAGAATACCAATTATAGCAATTGGCTTCAACAGATTG 60

OY 61 CTTGCTCATAGGATTCATGGCTGTGTTACATTAACTGATTCGTTCAACAGATTG 120  
 Db 61 CTTGCTCATAGGATTCATGGCTGTGTTACATTAACTGATTCGTTCAACAGATTG 120

OY 121 AATTTACACATTCACTTAAAGAAAGAGAGGAATGATGGGTTTGCTCG 180  
 Db 121 AATTTACACATTCACTTAAAGAAAGAGAGGAATGATGGGTTTGCTCG 180

OY 181 TTAACTTAAGTAGTCAGAAAGTTAGTCATGGTTACTGTGCTTATGTTGAA 240  
 Db 181 TTAACTTAAGTAGTCAGAAAGTTAGTCATGGTTACTGTGCTTATGTTGAA 240

OY 241 AATGAAAGGAACTTCAAACGGTTCTAGGGAAATGAGTTAGTCATCAT 300  
 Db 241 AATGAAAGGAACTTCAAACGGTTCTAGGGAAATGAGTTAGTCATCAT 300

OY 301 AATGTCCEGGCACTGTGACATAGATGTTTATTTTCAGTAGTTGCTCAT 360  
 Db 301 AATGTCCEGGCACTGTGACATAGATGTTTATTTTCAGTAGTTGCTCAT 360

OY 361 GTTTTGATTAACTTAACTCAATTCAATTCAAAACCTAGAQGACCAA 420  
 Db 361 GTTTTGATTAACTTAACTCAATTCAAAACCTAGAQGACCAA 420

OY 421 CAGCTCTCTCAATTGTAACAGAACAGAAGTTGGTGTGACCTAAAGACCTCC 480  
 Db 421 CAGCTCTCTCAATTGTAACAGAACAGAAGTTGGTGTGACCTAAAGACCTCC 480

OY 481 ATGAAAGCTCTGAGTCACCTGCCATTGCTCCATTAGCCATTCCA 540  
 Db 481 ATGAAAGCTCTGAGTCACCTGCCATTGCTCCATTAGCCATTCCA 540

OY 541 CCCGGCAACAAAGCTCGACCTAGGCCAACACTCACAATCTGCTCGCCGC 600  
 Db 541 CCCGGCAACAAAGCTCGACCTAGGCCAACACTCACAATCTGCTCGCCGC 600

OY 601 AATGGGCGACGCTCTCTCGACTCAATTCACTCCGAACTCCCTCC 660  
 Db 601 AATGGGCGACGCTCTCTCGACTCAATTCACTCCGAACTCCCTCC 660

OY 661 TGGCACACGCCAACACCCGCACTCTGTTCTGCAACATCTGATGCGRC 720  
 Db 661 TGGCACACGCCAACACCCGCACTCTGTTCTGCAACATCTGATGCGRC 720

OY 721 GAAGCCCAQGTCCCCATCCCATGATTCTACCAAGGTATTAGGAGCTCACACATTC 780  
 Db 721 GAAGCCCAQGTCCCCATCCCATGATTCTACCAAGGTATTAGGAGCTCACACATTC 780

OY 781 TTACCGATGGAATCAGAGAGCATTCGAACCTGAAACCCGAATTGGT 840  
 Db 781 TTACCGATGGAATCAGAGAGCATTCGAACCCGAATTGGT 840

OY 841 TTGGCGAGACCTTATCAGCGGAAGACAGATTCAGCTGCTGGAACCTCG 900  
 Db 841 TTGGCGAGACCTTATCAGCGGAAGACAGATTCAGCTGCTGGAACCTCG 900

OY 901 TCTATCCTCGGTAGAGAGTACATGATGAGAGCTCA 960  
 Db 901 TCTAATCCCTGGTAGAGAGTACATGAGGCTCTTGTGATGATGAGAGCTCA 960

OY 961 GTCATCACTGATGTTCTGGGATAGGTTATTCGATTCTGGAATATAAAGTTCTC 1020  
 Db 961 GTCATCACTGATGTTCTGGGATAGGTTATTCGATTCTGGAATATAAAGTTCTC 1020

OY 1021 GTTTAACTTCAAGATGGATAGGTTATTCGATTCTGGAATATAAAGTTCTC 1080  
 Db 1021 GTTTAACTTCAAGATGGATAGGTTATTCGATTCTGGAATATAAAGTTCTC 1080

OY 1081 TCTCTGTTATTCAGAGGGGTGAGATGAGATGTTCTGGGTGAGGCTTC 1140  
 Db 1081 TCTCTGTTATTCAGAGGGGTGAGATGAGATGTTCTGGGTGAGGCTTC 1140

OY 1141 GCTTAAGGGAGGGTGCCTAAGCTGTTAAGCAGATGTTGGTTTAGTTTGCGCTTC 1200  
 Db 1141 GCTTAAGGGAGGGTGCCTAAGCTGTTAAGCAGATGTTGGTTTAGTTTGCGCTTC 1200

OY 1201 GTTTCTCTGTTGAGGTTACCTGATTTATRACTGTTA 1260  
 Db 1201 GTTTCTCTGTTGAGGTTACCTGATTTATRACTGTTA 1260

OY 1261 TGAGTTCTGAGGAACCTTGAAGCTTACAGGTAGTTGACTCTCTGGTAATTG 1320  
 Db 1261 TGAGTTCTGAGGAACCTTGAAGCTTACAGGTAGTTGACTCTCTGGTAATTG 1320

OY 1321 ACAGGGTTGGCTTATAGAACCTTCTGATTTGAGTTGACTCTGGTGTGTA 1380  
 Db 1321 ACAGGGTTGGCTTATAGAACCTTCTGATTTGAGTTGACTCTGGTGTGTA 1380

OY 1381 GGAGGAAGGAGCAAGTAGCCCTGACCGGATTAGTGACAAATGATGAGACTTGG 1440  
 Db 1381 GGAGGAAGGAGCAAGTAGCCCTGACCGGATTAGTGACAAATGATGAGACTTGG 1440

OY 1441 AGAGATCACTCCGGTTATGCTGGAGCTACTTGCGTTACCGCTGGTGTGATGATTCG 1500  
 Db 1441 AGAGATCACTCCGGTTATGCTGGAGCTACTTGCGTTACCGCTGGTGTGATGATTCG 1500

OY 1501 TGCACAAAGCTAAATGGTTAACGGGTGGGGAAATTGTTGGCTGTGGAGGG 1560  
 Db 1501 TGCACAAAGCTAAATGGTTAACGGGTGGGGAAATTGTTGGCTGTGGAGGG 1560

OY 1561 TGGCAAAAGCTAAATGGTTAACGGGTGGGGAAATTGTTGGCTGTGGAGGG 1620  
 Db 1561 TGGCAAAAGCTAAATGGTTAACGGGTGGGGAAATTGTTGGCTGTGGAGGG 1620

OY 1621 ACCATGACCTGGCTGAGGAGTACAGTTAGAFACTTTTAACTTCTAGCA 1680  
 Db 1621 ATGATGAGGCTGCTGAGGAGTACAGTTAGAFACTTTTAACTTCTAGCA 1680

OY 1681 TGATATAACTTCTGTTCTGAGGAGTACAGTTAGAFACTTTTAACTTCTAGCA 1740  
 Db 1681 TGATATAACTTCTGTTCTGAGGAGTACAGTTAGAFACTTTTAACTTCTAGCA 1740

OY 1741 GCTACCCAAAGCAATTCAGAGAGTACAGTTAGAFACTTTTAACTTCTAGCA 1800  
 Db 1741 GCTACCCAAAGCAATTCAGAGAGTACAGTTAGAFACTTTTAACTTCTAGCA 1800

QY	1801	GTCGCTCAAGCTTATTGTTAGGAGGCCACATTTCAGGATTCAGGTTGATAACCAATT	1860	Db
Db	1861	CAGCACTCAGCAGGCTAGGTTATGGTATGGTATGGATCTGGATGTGTAATGAACTTC	1860	QY
QY	1861	CAGCACTCAGCAGGCTAGGTTATGGTATGGTATGGATCTGGATGTGTAATGAACTTC	1860	Db
Db	1921	CGGAAATTTGGAGAGTAGCTTCGGCTAGAAGGGACTCTCTGCACTGTTAGGATTC	1980	QY
QY	1921	CGGAAATTTGGAGAGTAGCTTCGGCTAGAAGGGACTCTCTGCACTGTTAGGATTC	1980	Db
QY	1981	AAAGTGTAATCCCGTTGCTGTTGGCTTGAGAATGACAGTGGATCAATAGGAT	2440	QY
QY	1981	AAAGTGTAATCCCGTTGCTGTTGGCTTGAGAATGACAGTGGATCAATAGGAT	2440	Db
Db	2041	CCAGCTATGTTGGAGTTGGCTGTTGGAGAATCAATCGTGTGACCAATGATGATG	2100	QY
QY	2041	CCAGCTATGTTGGAGTTGGCTGTTGGAGAATCAATCGTGTGACCAATGATGATG	2100	Db
Db	2101	GGACTATSCAATGTTGGAACCTGGTGGCAGGGGTTCTTCTAGGTCAGAAC	2160	QY
QY	2101	GGACTATSCAATGTTGGAACCTGGTGGCAGGGGTTCTTCTAGGTCAGAAC	2160	Db
QY	2161	ACCHAGATAAAATTAACTCGGGACTATAGTATGATCCTATGGTTGAGTAC	2220	QY
QY	2161	ACCHAGATAAAATTAACTCGGGACTATAGTATGATCCTATGGTTGAGTAC	2220	Db
Db	2221	TTCGAAGAGTGGGGAGTTCAGGTTCTCTTACGGTCTCTTACGGTCTCTTAC	2280	QY
QY	2221	TTCGAAGAGTGGGGAGTTCAGGTTCTCTTACGGTCTCTTACGGTCTCTTAC	2280	Db
Db	2221	TTGAAAGAGTGGGGAGTTCAGGTTCTCTTACGGTCTCTTACGGTCTCTTAC	2280	QY
QY	2221	TTGAAAGAGTGGGGAGTTCAGGTTCTCTTACGGTCTCTTACGGTCTCTTAC	2280	Db
Db	2281	ATTCGAGCGAGCATGTAAGGTAGTGTCTATGGAGGTTTCCTRCC	2340	QY
QY	2281	ATTCGAGCGAGCATGTAAGGTAGTGTCTATGGAGGACTCTGGCTTCCTRCC	2340	Db
QY	2341	CGCTATACGATAGAACTCGGCTAACCAAGGGTGCAGAGCAGTTAGTGA	2400	QY
QY	2341	CGCTATACGATAGAACTCGGCTAACCAAGGGTGCAGAGCAGTTAGTGA	2400	Db
Db	2401	GATCCCTGTTGTAACAACTTAGGCCGTGATGGCCCTGTTATGCAAGACT	2460	QY
QY	2401	GATCCCTGTTGTAACAACTTAGGCCGTGATGGCCCTGTTATGCAAGACT	2460	Db
Db	2401	GATCCCTGTTGTAACAACTTAGGCCGTGATGGCCCTGTTATGCAAGACT	2460	QY
QY	2461	GTAAGACCTCTGAAACTTGAACATAATGATGATTCGAAATTGGCTCTCAG	2520	QY
Db	2461	GTAAGACCTCTGAAACTTGAACATAATGATGATTCGAAATTGGCTCTCAG	2520	Db
QY	2521	AGTAGCGTTGATGAACACTCTGTTGAAATGTCCTGCTGATGTTAAAGGGCA	2580	QY
QY	2521	AGTAGCGTTGATGAACACTCTGTTGAAATGTCCTGCTGATGTTAAAGGGCA	2580	Db
Db	2581	GTGAGAGATCTAGCTGCTGCTGCTGCGCAATGGACTGTTACTGTTAGGAT	2640	QY
QY	2581	GTGAGAGATCTAGCTGCTGCTGCTGCGCAATGGACTGTTACTGTTAGGAT	2640	Db
Db	2641	TTCTTAAGCACCTCACTTTCACCAAGGATATGGTTCTCTCATGATCTGAT	2700	RESLT 3
Db	2641	TTCTTAAGCACCTCACTTTCACCAAGGATATGGTTCTCTCATGATCTGAT	2700	ADJ38210
QY	2701	GTGCTTACCATAGTTGATGTTAAATGATGCAATTTCATATCTGCAITGTC	2760	ID ADJ38210 standard; DNA: 2637 BP.
QY	2701	GTGCTTACCATAGTTGATGTTAAATGATGCAATTTCATATCTGCAITGTC	2760	XX
Db	2761	TGCTTGTGTTGAGCTAGAACAATAGTTCCTCACTTAATCATGTC	2820	AC ADJ38210;
QY	2761	TGCTTGTGTTGAGCTAGAACAATAGTTCCTCACTTAATCATGTC	2820	XX DT 06-MAY-2004 (first entry)
Db	2821	AAGATTAACAGTCTGAGTAACTTCACTAATGTCGAAATTGATCAA	2880	XX DE Plastid division-related Arc6 orthologue gene 3.
QY	2821	AAGATTAACAGTCTGAGTAACTTCACTAATGTCGAAATTGATCAA	2880	XX KW prokaryotic type; plastid division; Ptn2; ARC6; ARC5; Pzo; plant cell; herbi-
Db	2881	CTGAGACAGAAATGTAATTCACTCTAACATTCTCTTAAAGATAACGTTAGGAT	2940	XX OS Arabidopsis thaliana.
PN	W0200401003-A2.			



QY	ATGAACTGGCGTATGGGTGGCCTTAACTGAGTCAGTGAGGATCACATATAAGGAATCCAGCTA	2047	3068 GTATCCAAGPGCGAGAGTTAACPTCAGCTTGGGCTTGGGCTGATCACCCTAGAAATG	3127	
Db	ATGAACTGGCGTATGGGTGGCCTTAACTGAGTCAGTGAGGATCACATATAAGGAATCCAGCTA	1428	GTATCCAAGPGCGAGAGTTAACPTCAGCTTGGGCTTGGGCTGATCACCCTAGAAATG	2219	
QY	2048 TCTGTGAGTTGTTGGAAATTAACTCGTGTGACATGATGATGACATGATGATCTCCCTGACTAT	2107	TACCAAGGGTGGGAAATTACATCAATTCAATCAATTGTTGAGAACACTGTTGGACAT	3187	
Db	TCTGTGAGTTGTTGGAAATTAACTCGTGTGACATGATGATGACATGATGATCTCCCTGACTAT	1488	TACCAAGGGTGGGAAATTACATCAATTCAATCAATTGTTGAGAACACTGTTGGACAT	2226	
QY	2108 GCAAAATTGAAACCCGGTTGAGGGTTCTTCTTAGTGTAGAGACACAGAACAG 2167	2168 ATAAAATTAAACTCGGGACTATAGTGTGATCTANGTTGAGTTGACTCTGGAA 2227	GAAAGTTGGCTGAGAGTTAACPTCAGCTTGGGCTTGGGCTGATCACCCTAGAAATG	2248	
Db	ATAAAATTAAACTCGGGACTATAGTGTGATCTANGTTGAGTTGACTCTGGAA 2227	1549 ATAAAATTAAACTCGGGACTACTATGATGATCTATGGTTTGAGTACTTGAGAA	1608	GAAAGTTGGCTGAGAGTTAACPTCAGCTTGGGCTTGGGCTGATCACCCTAGAAATG	2308
QY	2228 GAGTGTGAGGTAGTCTCGGGTTCTCTTAGCTGCTGCTGCAACTATGCAAGGATGGAG 2287	1609 GAGTGTGAGGTAGTCTCGGGTTCTCTTAGCTGCTGCTGCAACTATGCAAGGATGGAG 1668	ACTGTGAAACTATCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	3367	
Db	GAGTGTGAGGTAGTCTCGGGTTCTCTTAGCTGCTGCTGCAACTATGCAAGGATGGAG 1668	2288 CGAGACATGTGAAAGCTAGTGTGATCGAGGACTCTGAGAAGCTGAGCTGCTGAGCTG	2347	3368 GGAAGCACTCTGGAGGAGTCGCTGCTGCTGCTGAGCTGAGCTGAGCTGAGCTG	3427
QY	1669 CGCAGCAGTGTGAAAGCTAGTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	1728	GGAAAGCACTCTGGAGGAGTCGCTGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	2428	
Db	GGAAAGCACTCTGGAGGAGTCGCTGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	2348 CAGATGAAACTCGGTGAAACCCAGGGATGTGAGAGACAGTTAGTGTAGATCCG	2407	3428 TACTGTGAGGACTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	3307
Db	2348 CAGATGAAACTCGGTGAAACCCAGGGATGTGAGAGACAGTTAGTGTAGATCCG	1729 CAGATGAAACTCGGTGCTGAGACCGAGGGATGTGAGAGACAGTTAGTGTAGATCCG	1788	3429 TACTGTGAGGACTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	2249
QY	2408 TCTGTGAACTGTGAGCTGCTGATGTCAGTGTGAGCTGAGCTGAGCTGAGCTG	2467	3488 ATACATGAGGCTCTGCTGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	3547	
Db	TCTGTGAACTGTGAGCTGCTGATGTCAGTGTGAGCTGAGCTGAGCTGAGCTG	1789 CCTCTGAAACTTGTGAACTTAATGATGAACTTGTGAGCTGAGCTGAGCTGAGCTG	1848	3489 ATACATGAGGCTCTGCTGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTG	2548
QY	2468 CCTCTGAAACTTGTGAACTTAATGATGAACTTGTGAGCTGAGCTGAGCTGAGCTG	2527	3548 CGAGATCTCTTGTGCTGAAATCTCTCTCTAAGTAGTTAGTGTATAATGACAA	3607	
Db	CCTCTGAAACTTGTGAACTTAATGATGAACTTGTGAGCTGAGCTGAGCTGAGCTG	1849 CCTCTGAAACTTGTGAACTTAATGATGAACTTGTGAGCTGAGCTGAGCTGAGCTG	1908	3549 CGAGATCTCTTGTGCTGAAATCTCTCTCTAAGTAGTTAGTGTATAATGACAA	2608
QY	2528 TTGATGAAACTCTGTGAAATGTCGGTGTGATATGTTAAGGGGCAAGTGAGA	2587	3608 AAAATTAACTTC 3620	QY	
Db	TTGATGAAACTCTGTGAAATGTCGGTGTGATATGTTAAGGGGCAAGTGAGA	1909 TTGATGAAACTACTCTGTTGAAATGTCGGTGTGATATGTTAAGGGGCAAGTGAGA	1968	2609 AAAATTAACTTC 2621	Db
QY	2588 TCTCTGCTGCTGGTGTGCAATTCGACTTCACTGTCACTGCTGAGTATTCTTA	2647	RESULT 4	QY	
Db	TCTCTGCTGCTGGTGTGCAATTCGACTTCACTGTCACTGCTGAGTATTCTTA	1969 TCCTAGCTGCTGGTGTGCAATTCGACTTCACTGTCACTGCTGAGTATTCTTA	2028	ADJ38129	ID ADJ38129 standard; cDNA; 2406 BP.
QY	2648 AAAGCAGCTCATCTTTCACGAGAGATGGGTCTCTCTCTCTCTCTCTCTCTCT	2707	XX	XX	XX
Db	AAAGCAGCTCATCTTTCACGAGAGATGGGTCTCTCTCTCTCTCTCTCTCTCT	2029 AAAGCAGCTCATCTTTCACGAGAGATGGGTCTCTCTCTCTCTCTCTCTCT	2088	ADJ38129;	ADJ38129;
QY	2708 CCCTAGGTATGATAATGCAATTCTGCATTCATGCTCAAATATGCTGT	2767	XX	XX	XX
Db	CCCTAGGTATGATAATGCAATTCTGCATTCATGCTCAAATATGCTGT	2089 CCA-----	2091	06-MAY-2004 (First entry)	06-MAY-2004 (First entry)
QY	2768 TTTGTGAGCTAAGAACATAGTCCACTTAATGATGCCAAAGTGTACCAAGATTA	2827	XX	XX	XX
Db	TTTGTGAGCTAAGAACATAGTCCACTTAATGATGCCAAAGTGTACCAAGATTA	2092 -----	2091	Arabidopsis thaliana Arc6-1 cDNA SeqID1.	Arabidopsis thaliana Arc6-1 cDNA SeqID1.
QY	2828 AGCAAGTGTGCTGACTTAATTCACTTCAATTATGCTGCTGAAATTGTGACA	2887	XX	XX	XX
Db	AGCAAGTGTGCTGACTTAATTCACTTCAATTATGCTGCTGAAATTGTGACA	2092 -----	2091	Arabidopsis thaliana.	Arabidopsis thaliana.
QY	2888 CAGAACTGTAATTCACTTCACTTCAACATTCTGTTAGATAACCTAGGATAGATGC	2947	XX	XX	XX
Db	CAGAACTGTAATTCACTTCACTTCAACATTCTGTTAGATAACCTAGGATAGATGC	2092 -----	2091	20-JUN-2003; 2003MO-US019536.	20-JUN-2003; 2003MO-US019536.
QY	3008 GTCTAGAGCTGACCTTCAGAACCTCCAGATGGATGCTAGCTGAGAGATA	3067	XX	XX	XX
Db	GTCTAGAGCTGACCTTCAGAACCTCCAGATGGATGCTAGCTGAGAGATA	2100 GTCTAGAGCTGACCTTCAGAACCTCCAGATGGATGCTAGCTGAGAGATA	2159	09-AUG-2002; 2002US-040224P.	09-AUG-2002; 2002US-040224P.
QY	2092 -----	2091	20-JUN-2003; 2003US-0600070.	20-JUN-2003; 2003US-0600070.	20-JUN-2003; 2003US-0600070.
Db	-----	-----	WO200401003-A2.	WO200401003-A2.	WO200401003-A2.
QY	-----	-----	31-DEC-2003.	31-DEC-2003.	31-DEC-2003.
Db	-----	-----	XX	XX	XX
QY	-----	-----	OS	OS	OS
Db	-----	-----	XX	XX	XX
QY	-----	-----	PN	PN	PN
Db	-----	-----	XX	XX	XX
QY	-----	-----	XX	XX	XX
Db	-----	-----	PA (UNMS ) UNIV MICHIGAN STATE.	PA (UNMS ) UNIV MICHIGAN STATE.	PA (UNMS ) UNIV MICHIGAN STATE.
QY	-----	-----	DR	DR	DR

XX  
PT New isolated Ptn2, ARCS and/or Pzo-like nucleic acid sequences, useful  
PT for further characterizing plastid division in plant cells, and in  
PT varying agronomic and horticultural characteristics of economically  
PT important plants.

XX  
PS Claim 1; SEQ ID NO 1; 287pp; English.

CC This invention relates to novel prokaryotic type or plastid division and  
related genes and proteins. In particular, the invention relates to novel  
CC Ptn2 (ARC6), ARC5 and Pzo-like genes and polypeptides. The methods and  
compositions of the present invention are useful for further  
CC characterising plastid division in plant cells, in order to vary  
CC agronomic and horticultural characteristics of economically important  
CC plants, such as crop, ornamental and woody plants. They can also be used  
CC related to the invention.

XX  
SQ Sequence 2406 BP; 612 A; 492 C; 618 G; 684 T; 0 U; 0 Other;  
Query Match 47.0%; Score 1724; DB 12; Length 2406;  
Best Local Similarity 79.2%; Pred. No. 0;  
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;

QY 481 ATCGAAGCTGTGTCAGTCGACATGGCTCTTCCCATTCCATTAATGCCATTACCA 540  
1 ATCGAACGCTGTGACTCAGCTGGCATTTGCTTCCCATATTGCCATTACCA 60

QY 541 CGCGGAGACAAAGCTCGACCTPAGCCCAAACCTCTAACATATGCTCGCCGC 600  
61 CGCGGAGACAAAGCTCGACCTPAGCCCAAACCTCTAACATATGCTCGCCGC 120

QY 601 AATGGGGCGACCGCTCTCTCGACTCAATTACCTCCATTCTCTCTCTCTCC 660  
121 AATGGGCGACGCTCTCTCGACTCAATTACCTCCATTCTCTCTCTCC 180

QY 661 TGGCCGACCGCCACCCACCGCAGCTCTGCTCTGCGCACATCTTGTCGCC 720  
181 TTGCCCCACCGCCACCCACCGCAGCTCTGCTCTGCGCACATCTTGTCGCC 660

QY 721 GAAGGCCAGTCCCATCCCATGATTCTACAGGATTAGAGCTCAACACATT 780  
241 GAAGGCCAGTCCCATCCCATGATTCTACAGGATTAGAGCTCAACACATT 300

QY 781 TMAACCGATGGATCAGAAGAGCTTCGAGCTTAGGGTTGAAACCCGGCAATTGGT 840  
301 TTACCGATGGATCAGAAGAGCTTCGAGCTTAGGGTTGAAACCCGGCAATTGGT 360

QY 841 TTAGCGGAGACCTTTATCAGCGGAGACAACTCTCAAGCTGCGAACCTG 900  
361 TTACGGGAGACCTTTATCAGCGGAGACAACTCTCAAGCTGCGAACCTG 420

QY 901 TCAATCTCGGCTCTAGAGAGATGACATGAGGTCTCTCTATGATGAGAGCTCA 960  
421 TCAATCTCGGCTCTAGAGAGATGACATGAGGTCTCTCTATGATGAGAGCTCA 480

QY 961 GTCACTACTGATGTTCTCTGGATAGGTTATTGATTCGATAATAAGTTCTC 1020  
481 GTCACTACTGATGTTCTCTGGATAGGTTATTGATTCGATAATAAGTTCTC 504

QY 1021 GTTTAACTTCATGATGAGAGAGAACTTTATCAGTGAAGGTTCTGGAGAC 1080  
505 ----- -AAGGTTCTGGGC 518

QY 1081 TCTCTGTGATTCAGAAGAGGTTGAGACTGAGATAGTCTCTGGGTTGAGCTCT 1140  
519 TCTCTGTGATTCAGAAGGTTGAGACTGAGATAGTCTCTGGGTTGAGCTCT 578

QY 1141 GCTTAAGGAGAGCTGCTTAAAGCTGAGATGGTTAGGTTAGTATGGCTTG 1200  
579 GCTTAAGGAGAGCTGCTTAAAGCTGAGATGGTTAGGTTAGTATGGCTTG 638

QY 1201 GTTGCTCGATGTCGAGGGATCTATGGCATTGATCCACCTGATTATGAGTCA 1260

Db 639 GTTCTCGATGTCGAGGGATCTATGGCATTGAGTACCTGATTTACTGGTTA 698  
Db 1281 TGAGTTCTGAGGAACCTTGAGCTTACAGGTTAGTTGACTTGCTTGGTAATTTG 1320  
Db 639 TGAGTTCTGAGGAACCTTGAGCTTACAGGTTAGTTGACTTGCTTGGTAATTTG 730

QY 1321 ACGAGCTTGGCTTATAAGAAACTTCTGATTTGATTTGAGCTTGGCTA 1380  
Db 731 ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- 731

QY 1381 GGAGGAAGGAGCTAGCTTGCACCGATTAGCCTACACAATGAGACTTGG 1440  
Db 732 GGAGGAAGGAGCTAGCTTGCACGGATTACGGTACGGCACAATGAGACTTGG 791  
Db 1441 AGAGATCACTCCGGTTAGTCTCTGGCTTGAGGACTCTMGGCTTACGGCTGGTGTGATGATACG 1500  
QY 1441 AGAGATCACTCCGGTTAGTCTCTGGCTTGAGGACTCTMGGCTTACGGCTGGTGTGATGATACG 851

Db 912 TGGCATGAGCTCTGTTGGGGTTGACCCGGTGGAGAGTTATGATGAGGGTTT 1620  
QY 1561 TGGAGCATAGCTCTGTTGGGGTTGACCCGGTGGAGAGTTATGATGAGGGTTT 971  
Db 912 TGGCATGAGCTCTGTTGGGGTTGACCCGGTGGAGAGTTATGATGAGGGTTT 971  
QY 1621 AGGATGAGAGCTGCTGAGCAGGTTAGAGCTTACGTTAGATACCTTTAACTCTTAA 1680  
Db 972 AGGATGAGAGCTGCTGAGC----- ----- ----- ----- ----- 991

QY 1681 TGATATAACTTGTAGTTCTCATTTAATGATGTTGTTGCTGTTGATCTTTGTA 1740  
Db 992 ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- 1008  
Db 1741 GCTACCCGAGGATATTCAGGAGCTTCTGAGTTAGAATGAGCTTGACTTGCTT 1800  
Db 1009 GCTACCCGAGGATATTCAGGAGCTTCTGAGTTAGAATGAGCTTGACTTGCTT 1068  
Db 1801 GTGGCTCAAGCTTATGTTGAGAAGGCCACCTTTACAGGATGCTGATAGCAATT 1860  
Db 1069 GTCGCTCAAGCTTATGTTGAGAAGGCCACCTTTACAGGATGCTGATAGCAATT 1128  
QY 1861 CAGCAACTTCAGGGCTTAAAGCTGCTGAGTTCTGGGAGNTGTTGATGATAC 1920  
Db 1129 CAGCAACTTCAGGGCTTAAAGCTGCTGAGTTCTGGGAGNTGTTGATGATAC 1188  
Db 1921 CGGAAATAATCGGAGATAGACTCTGGCTCTGAGGCTGCTGACTGCTTATAGGC 1980  
Db 1189 CGGAAATAATCGGAGATAGACTCTGGCTCTGAGGCTGCTGACTGCTTATAGGC 1248  
Db 1981 AAAGTGTGAAATGCGCTGAGTGGCTTAGAGCTGAGGATTACAAATATAGGA 2040  
Db 1249 AAAGTGTGAAATGCGCTGAGTGGCTTAGAGCTGAGGATTACAAATATAGGA 1308  
QY 2041 CCAGCTATGTTGAGCTGGGATCAATGTTGAGCTGAGGATTACAAATATAGGA 2100  
Db 1309 CCAGCTATGTTGAGCTGGGATCAATGTTGAGCTGAGGATTACAAATATAGGA 1368

QY 2101 GGACTATGCAATGTTGAGCTGGGAGGCTGTTCTGAGGCTGCTTCTGAGGCT 2160  
Db 1369 GGACTATGCAATGTTGAGCTGGGAGGCTGTTCTGAGGCTGCTTCTGAGGCT 1428

QY 2161 ACCAAAGATAAAATTAACTCGGGGACTACTATGATGATCTATGTTGAGTAC 2220  
Db 1429 ACCAAAGATAAAATTAACTCGGGGACTACTATGATGATCTATGTTGAGTAC 1488

QY 2221 TTGGAAAGCTGGGAGGTTAGGTTCTGGCTTGAGCTGCTGCAACTATGGCAGG 2280  
Db 1489 TTGGAAAGCTGGGAGGTTAGGTTCTGGCTTGAGCTGCTGCAACTATGGCAGG 1548

QY 2221 ATGGAGCCGACATGAAAGCTAGCTGCTGAGGCTGAGCTGCTGAGGCT 2340



QY	601	AATGGGGCAGCGCTCTCTCGACTCAATTCACTCCGAACTCCCTCTCTCTCC	660	QY	1681	TGTATAACTTAGTTTCATTTATGTAATGTTGCTGCTAGGTGACCTTGTGA	1740	
Db	121	AATGGGGCAGCGCTCTCTCGACTCAATTCACTCCGAACTCCCTCTCTCC	180	Db	992	-----AGGTGATCTTTGTA	1008	
QY	661	TTGCCAACGCCAACACCCACGGCACTCTGTTCTCTGCCAACCATCTATTGATGRC	720	QY	1741	GCTACCCAGCAATATTCAGCAGAGTCATTGAAGTTAGAAGTGCACTGCT	1800	
Db	181	TTGCCAACGCCAACACCCACGGCACTCTGTTCTCTGCCAACCATCTATTGATGRC	240	Db	1009	GCTACCCAGCAATATTCAGCAGAGTCATTGAAGTTAGAAGTGCACTGCT	1068	
QY	721	GAAGGCCAGTCATCCATCCCATGATTCTACAGGTTAGAGCTAACACATT	780	QY	1801	GTCGCTAAGCTTATGTTAGAAGGCCAACCTTTACAGATGTAGCAAGTCAATT	1860	
Db	241	GAAGGCCAGTCATCCATCCCATGATTCTACAGGTTAGAGCTAACACATT	300	Db	1069	GTCGCTAAGCTTATGTTAGAAGGCCAACCTTTACAGATGTAGCAAGTCAATT	1128	
QY	781	TMAACCGATGGATCAGAGAGGATTGAGCTAGGGTTGAAACCCCCGAAATTG	840	QY	1881	CAGCAACTTCAGCAGGCTAAGTGTGCTATGGAGATTCTGGGAGTTCTGGGAGTGTGATGACA	1920	
Db	301	TTACCGATGGATCAGAGAGGATTGAGCTAGGGTTGAAACCCCCGAAATTG	360	Db	1129	CAGCAACTTCAGCAGGCTAAGTGTGCTATGGAGATTCTGGGAGTTCTGGGAGTGTGATGACA	1188	
QY	841	TTAGCGGAGCACCTTTATCACCGGAGAACATTCTCAACTCTGTTGGAACACTG	900	QY	1921	CGGAAATAATGGAGATRAGACTCGGCTTAGAAGAGGACTGTGCACTGTATAGG	1980	
Db	361	TTAGCGGAGCACCTTTATCACCGGAGAACATTCTCAACTCTGTTGGAACACTG	420	Db	1189	CGGAAATAATGGAGATRAGACTCGGCTTAGAAGAGGACTGTGCACTGTATAGG	1248	
QY	901	TCTATCTCTGCTGAGAGGTTAGAGAGGTTACATGAGGTCTCTGTTGAGAGCTA	960	QY	1881	AAAGTTGATGAAATGCCGATGTTGTTGGCTAGAAGGGACTGTGCACTGTATAGG	2040	
Db	421	TCTATCTCTGCTGAGAGGTTACATGAGGTCTCTGTTGAGAGCTA	480	Db	1249	AAAGTTGATGAAATGCCGATGTTGTTGGCTAGAAGGGACTGTGCACTGTATAGG	1308	
Db	961	GTCTACACTGATGTTCTGGATTAGGTTATTGATTCGAAATAAAGTCCTC	1020	QY	2041	CCAGCTATTTGCGATTGAGTGTGCTATGGAGATTCTGGGAGTTCTGGGAGTGTGATGACA	2100	
Db	481	GTCTACACTGATGTTCTGGATT-----	504	Db	1309	CCAGCTATTTGCGATTGAGTGTGCTATGGAGATTCTGGGAGTTCTGGGAGTGTGATGACA	1368	
QY	1021	GTTTAAATTCTCATGATTTGGATTAAGGGAGAATTTAATCTGTTGAGGTT	1080	QY	2101	GGACTATGCAATTTGTTGAACTCTGTTGCGAGGTTGTTCTAGGTTTACAGAC	2160	
Db	505	-----	-ANGTCTGGGC	518	Db	1369	GGACTATGCAATTTGTTGAGAACCTGGCTAGAAGGGACTGTGCACTGTATAGG	1428
QY	1081	TCTCTGTTATCAAGAGGTTGAGACTGAGATAGTTCTGGGTTGAGGCT	1140	QY	2161	ACCAAGATTAATTAATCTGGGACTACTATGATGATCTATGGTTGAGTAC	2220	
Db	519	TCTCTGTTATCAAGAGGTTGAGACTGAGATAGTTCTGGGTTGAGGCT	578	Db	1429	ACCAAGATTAATTAATCTGGGACTACTATGATGATCTATGGTTGAGTAC	1488	
Db	1141	GCTTAAGGAGAGTTGCTTAAGCAGATGTTGTTAGTTGAGCTGTC	1200	QY	2221	TTGGAAAGTGAGGTTGAGGTTGAGCTGTCAGGGTTCTCTTGTGCTGCACTGGGAG	2280	
Db	579	GCTTAAGGAGAGTTGCTTAAGCAGATGTTGTTAGTTGAGCTGTC	638	Db	1489	TTGGAAAGTGAGGTTGAGGTTGAGCTGTCAGGGTTCTCTTGTGCTGAGCTGAGCTGAG	1548	
QY	1201	GTCTCTGTTGTTGAGGATCTGAGGAGATCTGATGTTATCTGGTA	1260	QY	2281	ATTCGACGCCGACATGTTGAAAGCTAGTTGCTATGCAAGGTTGTTCTAGGTTTACAGAC	2340	
Db	639	GTCTCTGTTGAGGAGCTTGAGGTTACAGGTTGTTGAGCTGTC	698	Db	1549	ATTCGACGCCGACATGTTGAAAGCTAGTTGCTATGCAAGGTTGTTCTAGGTTTACAGAC	1608	
QY	1261	TGGTTGTTGAGGAGCTTGAGGTTACAGGTTGTTGAGCTGTC	1320	QY	2341	CGCTATACAGATAGAACTCTGGCTGAACCAGGATGTTGAGCTGTC	2400	
Db	699	TGGTTGTTGAGGAGCTTGAGCTTACAGGTTGTTGAGCTGTC	730	Db	1609	CGCTATACAGATAGAACTCTGGCTGAACCAGGATGTTGAGCTGTC	1668	
Db	1321	ACCGCGTTGGCTTATAAGAACTTCTGATGTTGAGCTTGTGTA	1380	QY	2401	GATCCTGTTGTTACAGTGTAGCCGTTGAGCTGTC	2460	
Db	731	-----	A 731	Db	1669	GATCCTGTTGTTACAGTGTAGCCGTTGAGCTGTC	1728	
QY	1381	GGGGAGGAGGAGTAGCTTGAGCTTGAGGAGTTAGTGAGCAATTGATGAGCTTGTGTA	1440	QY	2461	GTAGACCCCTGAAACACTTGAACTATGTTGAGCTGTC	2520	
Db	732	GGGGAGGAGGAGTAGCTTGAGCTTGAGGAGTTAGTGAGCAATTGATGAGCTTGTGTA	791	Db	1729	GTAGACCCCTGAAACACTTGAACTATGTTGAGCTGTC	1788	
QY	1441	AGGATCACTCCCGTTAGCTGGAGCTACTGGCTTACCCCTGAGGTTATGATGAGCTTGTGTA	1500	QY	2521	AGTAGCGTTGATGAACTACTGTTGAGCTGTC	2580	
Db	792	AGGATCACTCCCGTTAGCTGGAGCTACTGGCTTACCCCTGAGGTTATGATGAGCTTGTGTA	851	Db	1789	AGTAGCGTTGATGAACTACTGTTGAGCTGTC	1848	
QY	1501	TGGAAAGACTAAATGTTAAGGGGTTGGCTGAGCTTGTGAGGTT	1560	QY	2581	GTAGAGATCTGTTGAGCTGTC	2640	
Db	852	TGGAAAGACTAAATGTTAAGGGGTTGGCTGAGCTTGTGAGGTT	911	Db	1849	GTAGAGATCTGTTGAGCTGTC	1908	
QY	1561	TGGAAAGACTAAATGTTAAGGGGTTGGCTGAGCTTGTGAGGTT	1620	QY	2641	TTCTTAAAGGAGCTCACTTCAACCGAAGGATGTTCTCTATGAGCTGTC	2700	
Db	912	TGGAAAGACTAAATGTTAAGGGGTTGGCTGAGCTTGTGAGGTT	971	QY	1909	TTCTTAAAGGAGCTCACTTCAACCGAAGGATGTTCTCTATGAGCTGTC	1968	
QY	1621	ACCATGAGCAGCTCTGAGGAGCTATACAGTTAGATACCTTTTTAATCTTAGCA	1680	Db	2701	GTGCTACCTAGGATGTTGAGCTGTC	2760	
Db	972	ACCATGAGCAGCTCTGAGC-----	991	Db	1969	GTGCTACCA-----	1978	
QY	2761	TCCTGTTGAGCTAGAGCATAGTCCCACCTAATAGTCCCACAGTTGAC	2820					

Db	1979	-----	1978	XX
QY	2821	AACATTAACAAGTGTGAGTAATTCACTAAATTAGTCGTTGAATTGGATCAA	2880	PR 20-JUN-2002; 2002US-0390140P.
Db	1979	-----	1978	PR 09-AUG-2002; 2002US-0402242P.
QY	2881	CTCTAGACAGAAAATGTAATTCACTCTCACACATTCTGTTAGATAACAGTAGGATTG	2940	PR 20-JUN-2003; 2003US-00600070.
Db	1979	-----	1978	PA (UNMS ) UNIV MICHIGAN STATE.
QY	2941	AGTTGCCATTAGTGTTGGCCTTCCAACTTTTCTTGTATTTCGATTTTCTGATTT	3000	XX
Db	1979	-----	1979	PT Osteryoung KW, Vitha S, Koksharova OA, Gao H;
QY	3001	ACGGTCAGTOAGGCTGAGATTAGAAGACACTCCAGAATGCTAGGCTAGCGAG	3060	XX DR WPI; 2004-092486/08.
Db	1980	ACGGTCAGTOAGGCTGAGATTAGAAGACACTCCAGAATGCTAGGCTAGCGAG	3039	DR P-PSDB; ADJ38207.
QY	3061	GAATATAGTATCCAGTGGAGAGATAGTCTGCTGCTTGGCCCTGATACCGAT	3120	XX PT New isolated Ptn2, ARCS and/or Pzo-like nucleic acid sequences, useful for further characterizing plastid division in plant cells, and in varying agronomic and horticultural characteristics of economically important plants.
Db	2040	GAATATAGTATCCAGTGGAGAGATAGTCTGCTGCTTGGCCCTGATACCGAT	2099	PT
QY	3121	AGAAATGTTACCGAGGTGGAGGATAATCTCAATTCAATGATGTTGAAACGT	3180	PT
Db	2100	AGAAATGTTACCG-----	2113	PS Disclosure; Fig 8; 287pp; English.
QY	3181	TGGACATGATAATAGTCGGCTGTTGATCTGTTATAGGTTGGCC	3240	XX
Db	2114	-----	2128	CC This invention relates to novel prokaryotic type or plastid division and related genes and proteins. In particular, the invention relates to novel Ftn2 (ARCS), ARCS and Pzo-like genes and polypeptides. The methods and compositions of the present invention are useful for further characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is that of a gene which is related to the invention.
QY	3241	GAATGCTGAGATTTGGACTGACAGAGCTGAACTCGCGAGCTGGTGTGTTAG	3300	CC Sequence 2406 BP; 612 A; 493 C; 618 G; 683 T; 0 U; 0 Other;
Db	2129	GAATGCTGAGATTTGGACTGACAGAGCTGAACTCGCGAGCTGGTGTGTTAG	2188	CC Query Match 46.8%; Score 1717.6; DB 12; Length 2406; Best Local Similarity 79.1%; Pred. No. 0; Mismatches 0; Matches 2402; Conservative 4; Indels 632; Gaps 5;
QY	3301	ATTTATACATGTTGAACTATCCTGTTGAGCTTGTGACAGTCTGAGCTGCGACCTG	3360	CC
Db	2189	ATTTATACATGTTGAACTATCCTGTTGAGCTTGTGACAGTCTGAGCTGCGACCTG	2248	CC
QY	3361	CCTGGTGGAGGAGACTCTGGAGAGTCGCTGCTGTATGTTGATGTTGTCAGAAGA	3420	CC
Db	2249	CCTGGTGGAGGAGACTCTGGAGAGTCGCTGCTGTATGTTGATGTTGTCAGAAGA	2308	CC
QY	3421	ACATGCTACTGATGAGAACCTACACACAGATACCAAGTTCTGTCGAGTCAG	3480	CC
Db	2309	ACATGCTACTGATGAGAACCTACACACAGATACCAAGTTCTGTCGAGTCAG	2368	CC
QY	3481	GCTGGAATACTCACTGAAAGCTCTTCTGCATCAA	3518	CC
Db	2369	GCTGGAATACTCACTGAAAGCTCTTCTGCATCAA	2406	CC
RESULT 6				XX
ID ADJ38208	standard: DNA: 2406 BP.			PR 20-JUN-2002; 2002US-0390140P.
XX				PR 09-AUG-2002; 2002US-0402242P.
AC				PR 20-JUN-2003; 2003US-00600070.
XX				PA (UNMS ) UNIV MICHIGAN STATE.
DT 06-MAY-2004 (first entry)				XX
DB Plastid division-related Arc6 orthologue gene 2.				PT Osteryoung KW, Vitha S, Koksharova OA, Gao H;
XX prokaryotic type; Plastid division; Rtn2; ARC6; Pzo; Plant cell; KW agronomic; horticultural; crop plant; ornamental plant; woody plant; KW herbicide target; gene; ds.				XX DR WPI; 2004-092486/08.
OS Arabidopsis thaliana.				DR P-PSDB; ADJ38207.
XX				XX
PN WO2004001003-A2.				XX
XX 31-DEC-2003.				XX
PD 20-JUN-2003; 2003WO-US019535.				XX
PP				XX



QV	3241 GAATGCTGAGATTTGGACTGACAGAGCTGAGTCAGAGCAGTGGCGAGSCGAGCTGAGTCAGAGCAGTGGCGAGTTGGTTATG 3300
Db	2129 GAACTGTGAGATTTGGACTGACAGAGCAGTGGCGAGTTGGTTATG 2188
QY	3301 ATTATACIUTGTGAACATCTGTGAGCTGAGCTCAGGAGTGGACCGCG 3340
Db	2189 ATTATACIUTGTGAACATCTGTGAGCTGAGCTCAGGAGTGGACCGCG 2248
QY	3361 CTCGGTGAGAACACTCTGGAGGGCTCTGTGATGTTGATCCAGAA 3420
Db	2249 CTCGGTGAGAACACTCTGGAGGGCTCTGTGATGTTGATCCAGAA 2308
QY	3421 ACAATGCTACTGTGAGTCAGAACCTCACACAAGAATACAGGAGTTCTGTCAGTCAGCTGAGCTCAGAGTGGACCCCG 3480
Db	2309 ACAATGCTACTGTGAGTCAGAACCTCACACAAGAATACAGGAGTTCTGTCAGTCAGCTGAGCTCAGAGTGGACCCCG 2368
QY	3481 GGTCGAGAMATCACTGAGGCTCTGTCATCTAA 3518
Db	2369 GGTCGAGAMATCACTGAGGCTCTGTCATCTAA 2406
RESULT 7	
ID ADJ38212 /C	
ID ADJ38212 standard; cDNA; 561 BP.	
XX AC	ADJ38212;
XX DT	06-MAY-2004 (first entry)
XX DE	Plastid division-related Arc6 orthologue cDNA 5.
XX KW	prokaryotic type; plastid division; Ftn2; Arc6; Arc5; Fzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target; gene; ss.
XX OS	Arabidopsis thaliana.
XX PN	WO200401003-A2.
XX PD	31-DEC-2003.
XX PR	20-JUN-2003; 2003WO-US019536.
XX PR	20-JUN-2002; 2002US-0390140P.
XX PR	09-AUG-2002; 2002US-0402242P.
XX PR	20-JUN-2003; 2003US-00600070.
PA (UNMS ) UNIV MICHIGAN STATE.	
PT Ostryoyoung KW, Vitha S, Koksharova OA, Gao H;	
DR WPI; 2004-082486/08.	
XX PT New isolated Ptn2, ARCS and/or Fzo-like nucleic acid sequences, useful for further characterising plastid division in plant cells, and in varying agronomic and horticultural characteristics of economically important plants.	
PS Disclosure; Fig 8; 287pp; English.	
CC This invention relates to novel prokaryotic type or plastid division and related genes and proteins. In particular, the invention relates to novel Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and compositions of the present invention are useful for further characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is a cDNA sequence which is related to the invention.	
SQ Sequence 561 BP; 167 A; 136 C; 107 G; 148 T; 0 U; 3 other;	
Query Match 11.0%; Score 402.2; DB 12; Length 561;	
Best Local Similarity 82.2%; Pred. No. 6.2e-94; Matches 514; Conservative 0; Mis matches 20; Indels 91; Gaps 1; 1;	
Matches 514; Conservative 0; Mis matches 20; Indels 91; Gaps 1; 1;	
QY 2970 TTCTCTCTGATTTCTTGTGATTAGGGTAGTCAGTCAGGATTCAGAG 3029	
Db 534 TTCTCTATGGRANCTGATGTCATACATAGGGTAGTCAGTCAGGATTCAGAG 475	
QY 3030 CACTTCAGATGGATGTCAGGACTGAGAATAGTTCAGAGGAGATA 3089	
Db 474 CACTTCAGATGGATGTCAGGACTGAGAATAGTTCAGAGGAGATA 415	
QY 3090 AGTCCTGGCTTGGCCCTATACGCCATAGAAATGTCACAGAGGAAATAA 3149	
Db 414 AGTCCTGGCTTGGCCCTATACGCCATAGAAATGTCACAGAGGAAATAA 370	
QY 3150 TCTACATTCAATCAATGTCGAAACTGTGACATGATATAGTCTGTCGCCTGTT 3209	
Db 369 ----- 370	
QY 3210 TGATTCGTGTTATTAGGTTGGAAGGGAAATCTGAGATGGACTGACAGAGCA 3269	
Db 369 ----- 326	
QY 3270 GCTGAAGACTGGCAGCTGGTTGGTTATGTTACACTGTCAGTCAGTCAGAGCA 3329	
Db 325 GCTGAAGACTGGCAGCTGGTTGGTTATGTTACACTGTCAGTCAGTCAGAGCA 266	
QY 3330 AGTGTGACAGTCAGCATGAGTGAACCCGTCCTGCTGGAGAACACTCTGGAGGACTCT 3389	
Db 205 GCTGTGCTATGATGTTGTCATCCAGAAACATGCTAGTCAGTCAGACCTACAA 146	
Db 265 AGTGTGACAGTCAGCATGAGTGAACCCGTCCTGCTGGAGAACACTCTGGAGGACTCT 206	
QY 3390 CCTGGTCACTGTGATGGTCATCCAGAAACAGTCAGTCAGACCTACACA 3449	
Db 205 GCTGTGCTATGATGTTGTCATCCAGAAACATGCTAGTCAGTCAGACCTACAA 146	
Db 3450 ACAGATACAGGTTCTGTCAGTCAGGTTGGAAATCACTGAGGCTCTGTCCT 3509	
Db 145 ACAGATACAGGTTCTGTCAGTCAGGTTGGAAATCACTGAGGCTCTGTCCT 86	
QY 3510 GCATCAATAATACTCATATGTCAGCTGCTGAGCTGCGAGATTCTTCTCTGAA 3569	
Db 85 GCATCAATAATACTCATATGTCAGCTGCTGAGCTGCGAGATTCTTCTCTGAA 26	
QY 3570 ATTCTCTCTGAGTTAGTGTATT 3594	
Db 25 ATTCTCTCTGAGTTAGTGTATT 1	
RESULT 8	
ID ADJ38264	
ID ADJ38264 standard; cDNA; 631 BP.	
XX AC	ADJ38264;
XX DT	06-MAY-2004 (first entry)
XX DE	Plastid division-related Arc6 orthologue cDNA 51.
XX KW	prokaryotic type; plastid division; Ftn2; Arc6; Arc5; Fzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target; gene; ss.
XX OS	Prunus persica.
XX PN	WO200401003-A2.
XX PD	31-DEC-2003.
XX PR	20-JUN-2003; 2003WO-US019536.
PR 20-JUN-2002; 2002US-0390140P.	
PR 09-AUG-2002; 2002US-0402242P.	
PR 20-JUN-2003; 2003US-00600070.	

(UNMS ) UNIV MICHIGAN STATE.  
XX PA  
PT Ostryoyoung KW, Vitha S, Koksharova OA, Gao H;  
XX  
DR WPI; 2004-082486/08.  
XX  
PT New isolated Ftn2, Arc5 and/or Fzo-like nucleic acid sequences, useful  
PT for further characterizing plastid division in plant cells, and in  
PT varying agronomic and horticultural characteristics of economically  
PT important plants.  
XX  
PS Disclosure; Fig 8; 287PP; English.  
XX  
CC This invention relates to novel prokaryotic type or plastid division and  
CC related genes and proteins. In particular, the invention relates to novel  
CC Ftn2 (Arc6), Arc5 and Fzo-like genes and polypeptides. The methods and  
CC compositions of the present invention are useful for further  
CC characterising plastid division in plant cells, in order to vary  
CC agronomic and horticultural characteristics of economically important  
CC plants, such as crop, ornamental and woody plants. They can also be used  
CC as herbicide targets. The present sequence is a cDNA sequence which is  
CC related to the invention.  
XX  
SQ Sequence 631 BP; 174 A; 123 C; 155 G; 178 T; 0 U; 1 Other;

ADJ38215;  
06-MAY-2004 (first entry)  
Plastid division-related Arc6 orthologue cDNA. 8.  
prokaryotic type; plastid division; Fts2; ARC6; ARC5; Fzo; plant cell;  
agronomic; horticultural; crop plant; ornamental plant; woody plant;  
hericide target; gene; ss.  
*Medicago truncatula*.  
W02004001003-A2.  
31-DEC-2003.  
20-JUN-2003; 2003WO-US019536.  
20-JUN-2002; 2002US-0390140P.  
09-AUG-2002; 2002US-0402242P.  
20-JUN-2003; 2003US-00600070.  
(UNMS ) UNIV MICHIGAN STATE.

卷之三

ID	ADJ38215
XX	06-MAY-2004 (first entry)
XX	DT
XX	DE Plastid division-related Arc6 orthologue cDNA 8.
XX	KW prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX	KW herbicide target; gene; ss.
XX	OS <i>Medicago truncatula</i> .
PN	XX
PN	WO2004001003-A2.
XX	XX
PD	31-DEC-2003.
XX	XX
PR	20-JUN-2003; 2003WO-US019536.
XX	XX
PR	20-JUN-2002; 2002US-03901405.
PR	09-AUG-2002; 2002US-0402242P.
PR	20-JUN-2003; 2003US-00600070.
XX	XX
PA	(UNMNS ) UNIV MICHIGAN STATE.
PT	XX
PT	Osterryoung KW, Vitha S, Koksharova OA, Gao H;
PT	DR WPI; 2004-082486/08.
XX	XX
PT	New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful for further characterizing Plastid division in plant cells, and in compositions of the present invention are useful for further characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is a cDNA sequence which is related to the invention.
PS	PS Disclosure; Fig 8; 287pp; English.
XX	XX
CC	This invention relates to novel prokaryotic type or plastid division and related genes and proteins. In particular, the invention relates to novel Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and compositions of the present invention are useful for further characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is a cDNA sequence which is related to the invention.
CC	CC
CC	CC
CC	CC
SQ	Sequence 660 BP; 149 A; 178 C; 152 G; 181 T; 0 U; 0 Other;
Query Match	5.7%; Score 208.8; DB 12; Length 660;
Best Local Similarity	60.3%; Pred. No. 1_6e-43;
Matches	429; Conservative 0; Mismatches 207; Indels 76; gaps 2;
Qy	561 ACCTGGCCACAGCACCTCTACAATCTGTCGCCAGCAATGGCGAACGCTCTCT 620
Db	17 ACCTAACGGTCTCATTCTCGCCGTCGCCAACGATAATGCGGGAGCGCAT 76
Qy	621 CTGGGACTCAATTCTCCCAATTCTCCCTCTCTCTCTGCCCCGGCACCCAC 680
Db	77 TTCCGATTCAATTCTCCGGACACTCCCTCTCTCCACACCCTCCGAC 136
Qy	681 CGGCCATCTGCTCTGCTGCCACCATATTCATGTCGCCGAGGCCAGGCCATCC 740
Db	137 AGTCATCTC-----ACTCTCTTACCTCTCCATAGACGCCAGCTGTCTACTCCC 190
Qy	741 CATGTTCTACAGGTTAGGAGTCACACACATTCTAACGGATGGATCAGAAG 800
Db	191 TCTCGACCTGTCACAAATCTCGGCCGCGAAACGCAATTCTCGGGATGATGGATCGGAG 250
Qy	801 AGCATTCGAAGCTAGGGTTGAAACGCCCGAAATTOGGTTTCAGGACGAGGCTTAAT 860
Db	251 AGCTTATGAGCGAAATTCTCGAGGCCCTCTCAGTAGTCAGTTCAGTATGAGCTTGAT 310

QY 861 CAGCCGGAGAGAGATCTTCGAAGCTCTGGAACACTCTCTAACTCCCTGGCTAGAG 920  
 Db 311 TAGTGTCTCGTCAAGTCTTCAGCTCTGGAACACTCTCTAACTCCCTGGCTAGAG 370  
 QY 921 AGAGTACAATGAGGCTCTCTGAGATGAGAACTACAGTCATCATGATGTTCTG 980  
 XX 371 AGAGTATAATCAAGCTCGTCGAGCATGAGACG 405  
 QY 981 GGATAGGTTAATTGATTTGGAATAAAGTTCTCTGTTTAATTCTGATGATGG 1040  
 Db 406 -----AAGATGAGGATCTT 420  
 QY 1041 ATAAAGGAGGAACTTTATCTAGTCAGGTTCTCCTGGGCTCTCTGTTGAGAG 1100  
 Db 421 CCTTTCTACTGAAATCCCTTGCACAACTTCTCTGAGCTGCTGCTGAG 480  
 QY 1101 GTGGTAGAGCTGAGATAGTTCTCGGTGTTGAGCTCTGCTGTTAGAGGAGGTCCTA 1160  
 481 CTGGAGAGACGGAGTTCTCGGATTGGAGGGTTTACTGAGAGAGGGTACCGA 540  
 QY 1161 AGTCGTTAGGCCAGATGCTGTTAGTTAGTATGCGCTTCTGAGTCTGAGGG 1220  
 Db 541 AGATGTTTAAGCAAGATGTTGTTGCTATGCGCTTCATATCTGAGCTTCTAGGG 600  
 QY 1221 ATGCTATGGATTGCCACCTGATTATTACTGTTATGAGTTGTGA 1272  
 Db 601 ATGCTATGGATTGCCACCTGATTCTGTTGAGATGCTGGA 652  
 RESULT 10  
 ADJ38206 ADJ38206 standard; DNA; 2283 BP.  
 ID XX  
 AC XX  
 XX ADJ38206;  
 DT 0-5-MAY-2004 (first entry)  
 XX DE Plastid division-related Arc6 orthologue gene 1.  
 XX KW prokaryotic type; plastid division; Ptn2; ARC6; ARCS; Pzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target; gene; ds.  
 XX OS Oryza sativa.  
 XX PN WO200401003-A2.  
 PD 31-DEC-2003.  
 XX PP 20-JUN-2003; 2003WO-US019536.  
 PR 20-JUN-2002; 2002US-0390140P.  
 PR 09-AUG-2002; 2002US-0402242P.  
 PR 20-JUN-2003; 2003US-05600070.  
 XX PA (UNMS ) UNIV MICHIGAN STATE.  
 PT Ostryoyoung KW, Vitha S, Koksharova OA, Gao H;  
 DR WPI: 2004-082486/08.  
 DR P-PSDB; ADJ38205.  
 XX PT New isolated Ptn2, ARCS and/or Pzo-like nucleic acid sequences, useful for further characterizing plastid division in plant cells, and in varying agronomic and horticultural characteristics of economically important plants.  
 XX PS Disclosure; Fig 8; 287pp; English.  
 XX CC This invention relates to novel prokaryotic type or plastid division and related genes and proteins. In particular, the invention relates to novel Ptn2 (ARC6), ARCS and Pzo-like genes and polypeptides. The methods and compositions of the present invention are useful for further

CC characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is that of a gene which is related to the invention.

SQ Sequence 2283 BP; 551 A; 576 C; 592 G; 564 T; 0 U; 0 Other;

Query Match 5.3%; Score 194; DB 12; Length 2283;  
 Best Local Similarity 55.6%; Pred. No. 2.1e-39; Matches 514; Conservative 0; Mismatches 300; Indels 110; Gaps 3; Matches 514; Conservative 0; Mismatches 300; Indels 110; Gaps 3;

QY 1368 TGAGCTTCTGTTAGGAGGAGGAGCTGGCTGACCGATTACCTGCAAAATG 1427  
 Db 716 TCAAGCTCTGAGGAGATGAGCTGAGCAAGAAATCTCGACACTGATGCTTCAGATG 775  
 QY 1428 ATGAGACTTGGAAAGAGATACTCGCGTATGTTCTGAGCTACTTGCTTACCGCTG 1487  
 Db 775 ATGANACTCTGAGGAGAGATACACCTCGTGTGTTGAGCTTCTCCTCTCCTTATG 835  
 QY 1488 GTGATGATTAAGCTGCTGCGAAAAGACTTAAGTTAAGCGGTGTCGGAAATTGTTG 1547  
 Db 836 ACACAGAGCATCATAGAAGCGCCAGAAAGGGCTCAAGCTGCGKAAGAACATTG 955  
 QY 1548 CTGTTGGAGGAGGTGGAGCATCAGCTCTTGTGGGGTTGACCGTGAGGAGTTGA 1607  
 Db 895 GCGTTGGAGGAGGTATGTCACCTGTCAGTGGAGGATTTCTCGTGAAGCCTCATGA 955  
 QY 1608 ATGAGCGGTTTTACGAAATACAGCTGTCAGCAGGATAACAGTTAGTACACCTTTT 1667  
 Db 956 ACGAGGCTTTTGAGGAGAGCATCAATTGAGATG----- 993  
 QY 1668 AATTCTTTAGCATGATATAACTTTAGTTCTGATTAAATGATGTTGTTGAGTT 1727  
 Db 994 ----- 993  
 DT 994 ----- 993  
 QY 1728 TGACCTTTGTTAGCTACCCAAAGCAATTCTCAGCAGAGTCATTGAGTTTACAGT 1787  
 Db 994 -GATTCTTTCAAACACCGAAATGCAATTCTCTGATGGTTGAATTACATG 1052  
 QY 1788 TGCACTTGCTCTGCTGGCTCAAGCTTAAATGGTAAGAGAACACCTTACAGATG 1847  
 Db 1053 AGCACTGCACTGTCGTCAGAATTAAAGAACGGCACAAATTCTCATGAGTC 1112  
 1848 TGATAGCAATTCCRGCAACTTCAGCAGCTGAGCTTAAAGTGGCTATGGAGATTCCTGCGAT 1907  
 Db 1113 GGATGATCTTTGAGAACACTCCAGAAGTCACATAGCTCTCATATTAGCT----- 1164  
 QY 1908 GTGTTGATGATACACCGAATTATGGAGTAGACTTCGGCTCTAGAAAGGGACTCTGTC 1967  
 Db 1165 -----TGTGATGATGAGATGGACTCTGATGCTGAGCTTCAATTGCAAAAGGCATTCTGCTC 1208  
 QY 1968 ACTGCTTATAGGCAAAAGTTGATGATGAGCTGCTGTTGAGCTGAGTTGAGGATTC 2027  
 Db 1209 ATTGCTAGTGGAGTTGAGCTGAGATGCAAGATGTTGAGATGATGATGTC 1268  
 QY 2028 ACAATATAGGAACTCGCTATGGAGCTGTTGAGGAAATTCAA---ATCGTGTCA 2084  
 Db 1269 ACCATACAGGACCCAAATTCTAGAGTTATGTCAGCAACTCTAGGATCAAGA 1328  
 QY 2085 CAATGATGATCTCCCTGAGCTATGCAATGTTGAAACCTGGTGGGGGTGTC 2144  
 Db 1329 GAATGATCTCTCAGSCTGTCAGACGCTTGTGAGCTTACCTGATGTC 1388  
 QY 2145 TCCTAGGTTGAGACACAAAGATAAATTAATGAGGACTACTATGATGATC 2204  
 Db 1389 TCCTAGGTTGAGACACAAAGATAAATTAATGAGGACTACTATGATGATC 1448  
 QY 2205 TATGTTTGTGTTACTGCGAGCTGCTGAGGAGTTGTCAGGGTCTCCCTTAGTGTCTC 2264  
 Db 1449 AGAAGTTTAAGCTAGTGAAGAGGATGGAGGGTGGTGGCTCTCATTTGGCTC 1508  
 QY 2265 TGCACATATGCGAAGGATTGGAGC 2288

**Db** 1509 ||| ||||| ||||| ||||| 1532  
**RESULT 11**  
ACN48781  
ID ACN48781 standard; cDNA; 607 BP.  
XX  
ACN48781;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Cotton primed seed EST Clone ID: LIB3825-027-Q6-K6-H1, SEQ:3562.  
XX  
Cotton; Plant; EST; expressed sequence tag; transgenic plant; seed; variety DP50B; library LIB3825; molecular tag; molecular marker; genetic mapping; molecular mapping; seed germination; plant growth; plant quality; plant yield; plant breeding; tissue printing; SB.  
XX  
OS *Gossypium hirsutum*.  
XX  
PN US2004123340-A1.  
XX  
PD 24-JUN-2004.  
XX  
PR 12-DEC-2001; 2001US-00021323.  
XX  
14-DEC-2000; 2000US-0255619P.  
XX  
PA (DEIK/ ) DEIKMAN J.  
PA (FENG/ ) FENG P C C.  
PA (FINC/ ) FINCHER K L.  
PA (ZIEG/ ) ZIEGLER T E.  
XX  
PT Deikman J, Feng PCC, Fincher KL, Ziegler TE;  
XX  
DR WPI; 2004-479808/45.  
XX  
PT New isolated nucleic acid molecule that encodes a plant protein or its fragment, useful for isolating a variety of agronomically significant genes associated with plant growth, quality or yield, and as molecular tags to map genes.  
XX  
PS Claim 1; SEQ ID NO 3562; 34pp; English.  
XX  
The invention relates to 17880 cotton expressed sequence tags (ESTs; ACN45220-ACN63059). The ESTs were isolated from cDNA libraries generated from primed or non-primed seeds from variety DP50B, mature seeds from variety Coker 312, Bollwell 96 Field, and Androecium tissue, gynoecium tissue, developing fibres, carpel walls and septa from variety Nucotton33B. The invention also relates to substantially purified proteins or their fragments encoded by nucleic acid molecules of the invention, and to transformed plants having a nucleic acid construct comprising a nucleic acid of the invention. The cotton ESTs are useful as molecular tags to isolate genetic regions, to isolate genes, to map genes, to determine gene function and to determine whether genes are members of a particular gene family. The nucleic acid molecule may be used for isolating a variety of agronomically significant genes associated with plant growth, quality, yield, and could also serve as links in metabolic and catabolic pathways. The nucleic acid molecules are also useful for identifying genes important in initiating and maintaining seed germination or that may be used to mitigate stresses encountered during seed germination. The ESTs additionally enable the acquisition of promoters and cis-regulatory elements which will be useful to express agronomically significant genes in these tissues and/or other tissues, and also permit the acquisition of molecular markers useful in breeding schemes, genetic and molecular mapping, and in cloning or agronomically significant genes. The nucleic acid molecules are further useful for detecting the expression level or pattern of a protein or mRNA and for detecting the presence or quantity of a protein by tissue printing. The present sequence represents a specifically claimed EST isolated from a cotton variety DP50B primed seed cDNA library (LIB3825). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340

**CC** was obtained in electronic format directly from the US patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340  
**Sequence** 607 BP; 172 A; 103 C; 164 G; 167 T; 0 U; 1 Other;

**SQ** Sequence 607 BP; 172 A; 103 C; 164 G; 167 T; 0 U; 1 Other;

**Query Match** Best Local Similarity 4.9%; Score 178.6; DB 13; Length 607; Matches 266; Conservative 0; Mismatches 100; Indels 18; Gaps 1;

**QY** 1986 TGATGAGTCGGCTATGTTGGGGCTTAGACGTGGAGATCACATATGAAATCCAGC 2045  
**Db** 5 TGAAGAGTGCCTCGTGGGGCTTAGACGTGGAGATCACATATGAAATCCAGC 64  
**QY** 2046 TATGTGGAGTGTGTTGGAGATCAATGTGATGACATGATGATCTCCCTGACT 2105  
**Db** 65 TATGTGAGATTGTCTTGAAACTCAGAACATGAGCATGAGCATCTCCGGGT 124  
**QY** 2106 ATGCAAATGTTGAGAACCTGGTGGAGGGGTGTTCTCTAGGTCAAGACACAA 2165  
**Db** 125 TTGCAACTGCTGGAGGATGCTAATGGAGTGTTCTAGATGATGATGCTTCTAGTGGAGACACAA 184  
**QY** 2166 AGATAAAAATTAAACTCGSGACTACTATGATGATGCTATGTTGAGTTCTAGTGGAG 2225  
**Db** 185 AGATATACATCAAGCTGGAGATTATGATGATGATGCTTCTAGTGGAGTTCTAGTGGAG 244  
**QY** 2226 AAGAGTGAGGAGTGTGAGGTCTCTTACTGCTGCTGCTCAACTATGGCAAGGATGG 2285  
**Db** 245 AAGGCTTGAGGAGCAGGTGGTCACCTTCTCTGCAGCAGCTATAGTGGAGTAGG 304  
**QY** 2286 AGCCGAG-----CATGTTGAAGCTAGTGCTATGAGGACTGAGA 2327  
**Db** 305 TCTGAGGCTACTGCACTCTGTATGATGATGTAAGGCTAGTCAATTGAGCATTGAGA 364  
**QY** 2328 AGTTTTCCTCCCGCTATACAGA 2351  
**Db** 365 GGTTGTTCCCTCTCGTGTCTGAGA 388

**RESULT 12**  
ADP93143  
ID ADP93143 standard; cDNA; 439 BP.  
XX  
AC ADP93143;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
DE Cotton expressed sequence tag, EST, #2154.  
XX  
KW Cotton; ss; EST; expressed sequence tag; Plant; plant protection; plant improvement; marker-assisted breeding.

**CC** *Gossypium hirsutum*; variety Nucotton33B.  
XX  
OS *Gossypium hirsutum*; variety Nucotton33B.  
XX  
PN US2004123338-A1.  
XX  
PD 24-JUN-2004.  
XX  
PF 08-DEC-2000; 2000US-00732627.  
XX  
PR 10-DEC-1999; 99US 0170355P.  
XX  
PA (FINC/ ) FINCHER K L.  
XX  
PI Fincher KL;  
XX  
DR WPI; 2004-479807/45.  
XX  
PT New substantially purified nucleic acid molecule that encodes a cotton protein or its fragment, useful as molecular tool for the targeting and isolation of novel genes for plant protection and improvement.  
XX  
PS Claim 1; SEQ ID NO 2154; 30pp; English.

Page 17

The invention relates to a substantially purified nucleic acid molecule that encodes a cotton protein or its fragment comprising an EST (expressed sequence tag) appearing as Adp0991-Adp5919. Also included are a substantially purified cotton protein or its fragment encoded by a nucleic acid molecule above and a transformed plant (having a nucleic acid molecule which comprises: an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; a structural nucleic acid molecule comprising one of the ESTs or their complements; a 3', non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as molecular tool for the targeting and isolation of novel genes for plant protection and improvement. The ESTs are useful for developing new strategies for understanding critical plant developmental and metabolic pathways, for isolating genes and promoters, for identifying and mapping the genes involved in developmental and metabolic pathways, and for determining gene function. The cotton nucleic acid molecules are useful as molecular tags to isolate genetic regions, isolate genes, map genes, and determine gene function. The nucleic acid molecules are useful for determining if genes are members of a particular gene family and for use in marker-assisted breeding programs. The present sequence is one of the 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed in the specification but are available in electronic format from the USPTO at seqdata.uspto.gov/sequence.html?docid=20040123338.

Matches 278; Conservative 0; Mismatches 141; Indels 9; Gaps 1

12 GTTAGATCGGCTGTGACTCCGACACTTGCTCCTTATTCGGACCTGTTAACCTCC 71

559 CGACGTAGCCACAAACACTCTACAACTATCTGGTCCGCCAGGAAATGGGGCAGCGCTT 618

619 CTCTCCGACTTCATTTCACCTCCG-----ATTCTCCCTCTCTCGCCACC 669

b  
132 CTGGCTGAACTCCAAATTCTCCCTGGCCCCGATTAATTCGGTCTCTCTCTCTCTCA 191

192 ACCGGCACTCTTCTCCCCCTTACCCCTCCCTTCCTCGCCCTTCCTCCGGAAACGGCAC 251

```

730 GTCCCCATCCCCATGATTCTACCGAGTATTGGAGCTAACACATTCTTAACGGAT 789
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |
252 GGGGGGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 311

```

790 GGAATCAGAAGAGCATTCGAAGCTAGGTTCCAAACGCCCAATCGTTTCAGGCAC 849

b  
312 GGAATTAGAGAGGCCCTATGAAGCCAAGGGTTCCAAAACCGCCCTCAATATGGTTCACTCAA 371  
850 GACGCTTAAATCAGCCGAGAGACAGATTCCTCAAGCTGCTTGGAACACTCTCTTAATCT 909

b 372 GACACCATATTAGCCGAGAACAGATTCTCTAGTGCTGTGAACCTTAACCT 431

b 916 CGGTCTAG 917  
432 GGCTCTAG 439

RESULT 13

X D ADJ38255 standard; cDNA; 545 BP.

X  
T  
06-MAY-2004 (first entry)

X 06-UNI-2004 LIBRARY  
E Plastid division-related Arc6 orthologue cDNA 42.

XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;  
 KW agromonic; horticultural; crop plant; ornamental plant; woody plant;  
 KW herbicide target; gene; ss.  
 XX  
 OS Gossypium arboreum.  
 XX  
 PN WO2004001003-A2.  
 XX  
 PD 31-DEC-2003.  
 XX  
 PR 20-JUN-2003; 2003WO-US019536.  
 XX  
 PR 20-JUN-2003; 2002US-0402242P.  
 PR 09-AUG-2002; 2002US-0402242P.  
 PR 20-JUN-2003; 2003US-00600070.  
 XX  
 PA (UNMS ) UNIV MICHIGAN STATE.  
 XX  
 PT Ostertagyoung KW, Vitha S, Kokharova OA, Gao H;  
 XX  
 DR WPI; 2004-082486/0B.  
 XX  
 PR New isolated Fm2, ARC5 and/or Fzo-like nucleic acid sequences, useful  
 for further characterizing plastid division in plant cells, and in  
 varying agronomic and horticultural characteristics of economically  
 important plants.  
 XX  
 PS Disclosure; Fig 8; 287pp; English.  
 XX  
 CC This invention relates to novel prokaryotic type or plastid division and  
 related genes and proteins. In particular, the invention relates to novel  
 Ftn2 (ARC5), ARC5 and Fzo-like genes and polypeptides. The methods and  
 compositions of the present invention are useful for further  
 characterising plastid division in plant cells, in order to vary  
 agronomic and horticultural characteristics of economically important  
 plants, such as crop, ornamental and woody plants. They can also be used  
 as herbicide targets. The present sequence is a cDNA sequence which is  
 CC related to the invention.  
 XX  
 SQ Sequence 545 BP; 164 A; 107 C; 120 G; 151 T; 0 U; 3 Other;  
 Query Match 4,4%; Score 161; DB 12; Length 545;  
 Best Local Similarity 69.6%; Pred. No. 4e-11; Matches 218; Conservative 0; Mismatches 95; Indels 0; Gaps 0;  
 Oy 3224 ATAGGTTCATGGCGGAATGCTGAGATTGGACTGACAGAGCTGAACTCGGCA 3223  
 Db 82 AGAGGTCTGATGGTCAAATGGAGACATGGAGAGCTGTCGA 141  
 Oy 3284 GCTTGGGTGGTTATGATTTACACTGTGAACTATCTGTGACAGTGACAGTC 3343  
 Db 142 GCTTGGGTGGTTATGATTTACACTGTGACAGTGACAGTC 201  
 Oy 3344 ACCAGATGGAAACCGGTCTCTGGTGGAGAAGCACTCTGGAGAGCTGTCTGTCTATCTGA 3403  
 Db 202 ACTAGTGGCCAGGAGCTGAGCTACTCTGGAGAATCCACCTGCTGACGTA 261  
 Oy 3404 TTGGTCATCCGAGAACATGCTGAGCTGAGCTACCTGGAGAATCACAAGATAGAATGAGT 3463  
 Db 262 CCTCTTCATCGGAGAACATGCTGAGCTACCTGAGGATCTGAGGATCTGCTACAACTTAACTATG 321  
 Oy 3524 TCATATGGTAGCAT 3536  
 Db 382 ATGTATAAAGCAT 394

ID	ACN48855; standard; cDNA; 552 BP.	XX
AC		XX
XX		XX
DT		DT
DE	Cotton primed seed EST Clone ID: LIB3825-027-Q6-N6-H1, SEQ:3636.	XX
XX	cotton; plant; EST; expressed sequence tag; transgenic plant; Seed; variety DP50B; library LIB3825; molecular tag; molecular marker; gene mapping; molecular mapping; seed germination; plant growth; plant quality; plant yield; plant breeding; tissue printing; ss.	KW
XX	Gossypium hirsutum.	KW
XX	US2004123340-A1.	PN
XX	24-JUN-2004.	PD
XX	12-DEC-2001; 2001US-00021323.	PP
XX	14-DEC-2000; 2000US-02555619P.	PR
PA	(DEIK/), FENG P C C., FINCHER K L., ZIEGLER T E.	PA
PA	(FINC/), FINCHER K L., ZIEGLER T E.	PA
XX	Deikman J, Feng PCC, Fincher KL, Ziegler TE;	PI
XX	WPI; 2004-473808/45.	XX
PT	New isolated nucleic acid molecule that encodes a plant protein or its fragment, useful for isolating a variety of agronomically significant genes associated with plant growth, quality or yield, and as molecular tags to map genes.	PT
PT	Claim 1; SEQ ID NO 3636; 34BP; English.	PS
XX	The invention relates to 1780 cotton expressed sequence tags (ESTs; ACN45220-ACN03099). The ESTs were isolated from cDNA libraries generated from primed or non-primed seeds from variety DP50B, mature seeds from variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium tissue, developing fibres, carpel walls and septa from variety Nuccton33B. The invention also relates to substantially purified proteins or their fragments encoded by nucleic acid molecules of the invention, and to transformed plants having a nucleic acid construct comprising a nucleic acid of the invention. The cotton ESTs are useful as molecular tags to isolate genetic regions, to isolate genes, to map genes, to determine gene function and to determine whether genes are members of a particular gene family. The nucleic acid molecule may be used for isolating a variety of agronomically significant genes associated with plant growth, quality, Yield, and could also serve as links in metabolic and catabolic pathways. The nucleic acid molecules are also useful for identifying genes important in initiating and maintaining seed germination or that may be used to mitigate stresses encountered during seed germination. The ESTs additionally enable the acquisition of promoters and cis-regulatory elements which will be useful to express agronomically significant genes in these tissues and/or other tissues, and also permits the acquisition of molecular markers useful in breeding schemes, genetic and molecular mapping, and in cloning of agronomically significant genes. The nucleic acid molecules are further useful for detecting the expression level or pattern of a protein or mRNA and for detecting the presence or quantity of a protein by tissue printing. The present sequence represents a specifically claimed EST isolated from a cotton variety DP50B seed cDNA library (LIB3825). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at seqdata.uspto.gov/sequence.html?docID=US2004123340	XX
XX	Sequence 552 BP; 153 A; 122 C; 110 G; 167 T; 0 U; 0 other;	SQ
Matches	217, Conservative 69 3%; Pred. No. 1.1e-30;	Best Local Similarity 0, Mismatches 96, Indels 0, Gaps 0;
OY	3224 ATAGCTTTCGATGGCGATGCTGAAGATTGGACTGAGCAGAGCTGCGCA	3283
Db	444 AGAGGTCTCGATGGTCAGTTGAGACATGGACAGATGGCGCAATGCTCA	385
QY	3284 GCTCTGGTTGTTATGATTTATGTTACACTGTTGAACTATCTGTTGAGCAGTGACGTC	3343
Db	384 GCTCTGGTTGTTATGATTTATGTTACACTGTTGAACTATCTGTTGAGCAGTGACGTC	325
QY	3344 NGCAGATGGAAACCCTGCTGTTGGAAAGAACCTGAGGAGCTGCTGCTGCTAATCTGA	3403
Db	324 ACTAGATGGCAGGGAGCTGTTGAGTGGAACTCTGAGTGGAGATCCACCTGCTGA	265
QY	3404 TTTCGTTCACTGAGCTGGAAATGACTGAAAGCTGCTGCTGCTGCTAATCTGA	3463
Db	264 TGTTCATCATCCGGAGAACATGCTCTAATGTAACCTCTACACCAGATAATGAGAT	205
QY	3464 TTTCGTTCACTGAGCTGGAAATGACTGAAAGCTGCTGCTGCTAATCTGA	3523
Db	204 GTCTTGTCTCACTCAGGCTGAAATCACTGAGGATCTGCTACAAATCTTAGCTAG	145
QY	3524 TCATATGTTGAGAT 3536	Db
Db	144 ATGTTAAAGCT 132	Db
RESULT 15	ADJ38217	Best Local Similarity 0, Mismatches 96, Indels 0, Gaps 0;
ID	ADJ38217 standard; cDNA; 608 BP.	XX
AC	ADJ38217,	XX
DT	06-MAY-2004 (first entry)	DT
XX	Plastid division-related Arc6 orthologue cDNA 10.	DE
XX	prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target; gene; ss.	KW
XX	Solanum tuberosum.	CC
PN	WO200401003-A2.	CC
XX	31-DEC-2003.	PD
XX	20-JUN-2003; 2003WO-US019536.	PP
XX	20-JUN-2002; 2002US-0390140P.	PR
PR	09-AUG-2002; 2002US-0402242P.	PR
PR	20-JUN-2003; 2003US-00600070.	PR
XX	(UNMS ) UNIV MICHIGAN STATE.	PA
XX	Osteryoung KW, Vitha S, Koksharova OH, Gao H;	PI
XX	WPI; 2004-082486/08.	XX
XX	New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful for further characterizing plastid division in plant cells, and in varying agronomic and horticultural characteristics of economically important plants.	PT
XX	This invention relates to novel prokaryotic type or plastid division and related genes and proteins. In particular, the invention relates to novel Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and compositions of the present invention are useful for further characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important	CC

CC plants, such as crop, ornamental and woody plants. They can also be used  
 CC as herbicide targets. The present sequence is a cDNA sequence which is  
 CC related to the invention.

XX Sequence 608 BP; 146 A; 167 C; 128 G; 167 T; 0 U; 0 Other;  
 SQ Query Match 4.3%; Score 158; DB 12; Length 608;  
 Best Local Similarity 63.2%; Pred. No. 2.6e-30;  
 Matches 264; Conservative 0; Mismatches 145; Indels 9; Gaps 1;

OY 581 CAACTATCGTCGCCAGCAAATGGCGGACCGCTCTCGACTCAATT----- 635  
 Db 169 CTAACTCTCCCGCTCTAGTGGCGATCTCTTCCTGCCTTGCGATTTCCTTC 228  
 Qy 636 ---CACCTCGATTCCTCTCTCTCTCTCCACCCACACCCGCGACTCCTCG 691  
 Db 229 CTTCACCAACCCCTCCGACTCATCGATTCTCGAATTCAACTCTACACTCCGTA 288  
 Qy 692 TCTCTCTGCCACCACTATGATCTCCGAACCCACGCTCCCATCCATGATTC 751  
 Db 289 CGACTATTCCTCCCTCTGTGCTCTTCAAGACCAACCACTTAAATGCTATGACTTT 348  
 Qy 752 ACCAGGTATTAGGAGCTAACACATTCTTAACCGATGGATACTAGAAGAGCATCGAG 811  
 Db 349 ATAGAGTCTCTGCTCTGAGCTTCACTCTCTCGSTGACGCTTATTGGAGTCTACGTG 408  
 Qy 812 CTGGGTTGAGAACCGGGGAAATCGGTTTGGGACGCGTTAACCGGGGAGAC 871  
 Db 409 CTAGAATTACAAACCTCCGACTAGGATACACTCAGGAAGGATGTTGGCGACGCC 468  
 Qy 872 AGATCTCGAAGCTCTCGGAACCTCTCTCTAACCTCGGTCTAGAAGAGCTACATG 931  
 Db 469 AGATCTCGAAGCTCTGTGAAACCTCTGCTACCTCTACCTCTGTAGAGATACATC 528  
 Qy 932 AACGCTCTCTGTGATGAGAAGACTACAGTCATCACTGATGTTCTGGATTAAGT 989  
 Db 529 AACGCTCTCTGTGATGAGAAGACTACAGTCATCACTGATGTTCTGGATTAAGT 586

Search completed: June 9, 2005, 17:27:57  
 Job time : 1826 secs

*THIS PAGE BLANK (USPTO)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## Om protein - protein search, using sw model

Run on: June 9, 2005, 17:28:08 ; Search time 94 Seconds  
(without alignments)  
329.692 Million cell updates/sec

Title: US-10-600-070-2

perfect score: 4063

Sequence: 1 MEALSHVGIGLSPFOLCRLP..... YEVFWNSKSGWKITEGSVLA S 801

Scoring table: BL0SUM62

Gapext 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_16Dec04:\*

1: geneseqDPI980s:\*

2: geneseqDPI990s:\*

3: geneseqDPI2000s:\*

4: geneseqDPI2010s:\*

5: geneseqDPI2020s:\*

6: geneseqDPI2003ab:\*

7: geneseqDPI2003bs:\*

8: geneseqDPI2004g:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	4063	100.0	801	ADJ8202
2	4051	99.7	801	ADJ8207
3	4051	99.7	801	ADJ8211
4	4051	99.7	801	ADJ8209
5	1775.5	43.7	760	ADJ8205
6	1646	40.5	324	ADJ8203
7	524.5	12.9	760	ADJ8247
8	498.5	12.3	789	ADJ8274
9	497	12.2	798	ADJ8244
10	497	12.2	798	ADJ8245
11	442	10.9	631	ADJ8242
12	442	10.9	631	ADJ8132
13	442	10.9	631	ADJ8240
14	431	10.6	819	ADJ8253
15	431	10.6	819	ADJ8251
16	429	10.6	714	ADJ8249
17	429	10.6	714	ADJ8250
18	423	10.4	673	ADJ8272
19	413	10.2	835	ADJ8270
20	401	9.9	716	ADJ8239
21	344.5	8.5	702	ADJ8235
22	320	7.9	651	ADJ8237
23	145.5	3.6	6095	AAU10701 Amino acid
24	143	3.5	2089	AAW08333 Cyclotide
25	139.5	3.4	954	AAW72752 DNA polym

## ALIGNMENTS

26	139.5	3.4	954	2 ADC87484 Confectio
27	138	3.4	2021	7 ADJ70511 Human hea
28	133.5	3.3	6973	5 ADC26982 Sorghum
29	133	3.3	1273	4 ABG0013 Novel hum
30	132.5	3.3	876	8 ADM13791 Bacillus
31	131.5	3.2	2111	4 AAB66471 Protein e
32	130.5	3.2	876	6 ABR2335 Bacillus
33	130	3.2	1082	5 ABR2283 Human pol
34	128.5	3.2	875	2 AAR80137 B.stearot
35	128.5	3.2	876	2 AAR80143 B.stearot
36	128.5	3.2	876	2 AAR80142 B.stearot
37	128	3.2	9234	7 ADC26980 Sorghum
38	127.5	3.1	2472	4 AAU33568 Pseudomon
39	127	3.1	1241	3 AAY3337 Human i
40	127	3.1	1241	6 ADJ44856 Human int
41	127	3.1	1241	7 ABR2339 Abi4249 Human int
42	127	3.1	1241	8 ADR46654 Cancer-as
43	126.5	3.1	1169	5 ABP29745 Streptoco
44	126.5	3.1	1224	5 ABP26662 Streptoco
45	126.5	3.1	2482	7 ABO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87484 Confectio  
AdJ70511 Human hea  
AdC26982 Sorghum  
AbG0013 Novel hum  
AdM13791 Bacillus  
AAB66471 Protein e  
Abi4335 Bacillus  
Abp6283 Human pol  
Aar80137 B.stearot  
Aar80143 B.stearot  
Aar80142 B.stearot  
Add26980 Sorghum  
Aau33568 Pseudomon  
Aay3337 Human i  
AdJ44856 Human int  
Abi4249 Human int  
Ad46654 Cancer-as  
Abp29745 Streptoco  
Abp26662 Streptoco  
AbO68221 Pseudomon

Ado87

CC	related to the invention.	DE	Plastid division-related Arc6 protein orthologue 2.
XX		XX	prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;
SQ	Sequence 801 AA;	KW	aeronomic; horticultural; crop plant; ornamental plant; woody plant;
		KW	herbicide target.
Matches	Best Local Similarity 100.0%; Score 4063; DB 8; Length 801;	XX	Arabidopsis thaliana.
801	Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;	XX	
OY	1 MEALSHVGIGLSPFOLCRLLPATTKRRSHNTSTICSAASKWADRLLSDPNFTSDSSSS 60	OS	W02004001003-A2.
Db	1 MEALSHVGIGLSPFOLCRLLPATTKRRSHNTSTICSAASKWADRLLSDPNFTSDSSSS 60	PR	31-DEC-2003.
OY	61 FATAATTATLVSPPSIDIQRPERHVPIPDYQVLAQTHFLTGIRRAFEARSKPQFG 120	PR	20-JUN-2003; 2003WO-US019336.
Db	61 FATAATTATLVSPPSIDIQRPERHVPIPDYQVLAQTHFLTGIRRAFEARSKPQFG 120	PR	20-JUN-2002; 2002US-0390140P.
OY	121 FSDDALISRQILOQACETLSNPSRSRREYNGLIDDEBEATVTDVKPGALCVLQEG 180	PR	09-AUG-2002; 2002US-040242P.
Db	121 FSDDALISRQILOQACETLSNPSRSRREYNGLIDDEBEATVTDVKPGALCVLQEG 180	PR	20-JUN-2003; 2003US-0060070.
OY	181 GETEIVLRVGEALKERLKPFSKQDVVLVMAFLDVSRA DALDPDFITGEYFVEAL 240	DR	DR
Db	181 GETEIVLRVGEALKERLKPFSKQDVVLVMAFLDVSRA DALDPDFITGEYFVEAL 240	N-PSDB,	ADJ38208.
OY	241 KLIQEEGASSLAPDLRQIDETEETBETRIVYVLLGLGDDYAKRNLGSVRNLLWS 300	XX	(UNMIS ) UNITV MICHIGAN STATE.
Db	241 KLIQEEGASSLAPDLRQIDETEETBETRIVYVLLGLGDDYAKRNLGSVRNLLWS 300	XX	Ostryoung KW, Vitha S, Koksharova OA, Gao H;
OY	301 VGGGASALVGGTREKFMNEAFLRMTAEQDLYFATPSNITAESFEVYEVALVAQA 360	XX	PT
Db	301 VGGGASALVGGTREKFMNEAFLRMTAEQDLYFATPSNITAESFEVYEVALVAQA 360	XX	WPI; 2004-082486/08.
OY	301 VGGGASALVGGTREKFMNEAFLRMTAEQDLYFATPSNITAESFEVYEVALVAQA 360	XX	DR
Db	301 VGGGASALVGGTREKFMNEAFLRMTAEQDLYFATPSNITAESFEVYEVALVAQA 360	XX	N-PSDB,
OY	361 FIGKKPHILQDAKQFOQLQAKVMAPEPAMWJDTRNNWEIDPGLERGLCALLIGKDE 420	XX	ADJ38208.
Db	361 FIGKKPHILQDAKQFOQLQAKVMAPEPAMWJDTRNNWEIDPGLERGLCALLIGKDE 420	XX	
OY	421 CRMWLGUDSESDSYRNPAVLFENSLNSRHDNDLPGCLLTTWLAGVVFPERFDTKD 480	CC	New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
Db	421 CRMWLGUDSESDSYRNPAVLFENSLNSRHDNDLPGCLLTTWLAGVVFPERFDTKD 480	CC	for further characterization of plastid division in plant cells, and in
OY	481 KFLCLGDXDDPMU5YL SERVEVVOGSPLAAATMARGHVRVAMQALQKFPSTYD 540	CC	varying agronomic and horticultural characteristics of economically
Db	481 KFLCLGDXDDPMU5YL SERVEVVOGSPLAAATMARGHVRVAMQALQKFPSTYD 540	CC	important plants.
OY	541 RN5ABPKDVOQETTSFSDPVGNNTGDRGSRGVFAEAEPSENITENDYAIRAGVSESSVD 600	XX	CC
Db	541 RN5ABPKDVOQETTSFSDPVGNNTGDRGSRGVFAEAEPSENITENDYAIRAGVSESSVD 600	XX	characterising plastid division in plant cells, in order to vary
OY	601 ETTVEMSVADMLKEASVKAIIAGVAGLISLFSOKYFLKSSSSQRKVNSSNESDVTI 660	CC	agronomic and horticultural characteristics of economically important
Db	601 ETTVEMSVADMLKEASVKAIIAGVAGLISLFSOKYFLKSSSSQRKVNSSNESDVTI 660	CC	plants, such as crop, ornamental and woody plants. They can also be used
OY	661 GSTRADDSSEALPMDARTAENTISKWOKIKSLAFGDRRIEMPPEVLDGRMKIWTBAA 720	CC	as herbicide targets. The present sequence is that of a protein which is
Db	661 GSTRADDSSEALPMDARTAENTISKWOKIKSLAFGDRRIEMPPEVLDGRMKIWTBAA 720	CC	related to the invention.
OY	721 EIMQGLGLDYDTIILKLSDVTSA DGRALVIALESA CSDLVH PENNA TDV RYT 780	XX	XX
Db	721 EIMQGLGLDYDTIILKLSDVTSA DGRALVIALESA CSDLVH PENNA TDV RYT 780	XX	
OY	781 RYKVFWMSKGWKLTTEGSVLAS 801	XX	
Db	781 RYKVFWMSKGWKLTTEGSVLAS 801	XX	
RESULT 2	781 RYKVFWMSKGWKLTTEGSVLAS 801	XX	
ID	781 RYKVFWMSKGWKLTTEGSVLAS 801	XX	
AC	781 RYKVFWMSKGWKLTTEGSVLAS 801	XX	
DT	06-MAY-2004 (first entry)	XX	
OY	361 FIGKKPHILQDAKQFOQLQAKVMAPEPAMWJDTRNNWEIDPGLERGLCALLIGKDE 420	XX	

Db ||||| FIKKKPHILODAKQFOOLQOAKWMAMIPAMLYDRNWEIDFGLERGCALLIGKVE 420  
 CC 361 KERFLGDXYDDPMVLSYLVERVEVQGSPLAATAATMARIAGAHTVKASAMQALQKVFPSPYTD 480  
 CC 421 CRMWGLGDSEDSQYRNPAIVFLENNSNRDDNDLPGICKLETLWLAGVPRFRDKD 480  
 CC 421 CRMWGLGDSEDSQYRNPAIVFLENNSNRDDNDLPGICKLETLWLAGVPRFRDKD 480  
 CC as herbicide targets. The present sequence is that of a protein which is  
 CC related to the invention.  
 XX SQ Sequence 801 AA;  
 Query Match 99.7%; Score 4051; DB 8; Length 801;  
 Best Local Similarity 99.8%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;  
 Matches 799; Conservative 0;  
 Oy 1 MERALSHVGIIGLSPPFOLCRPPLPATRKRSRNTTCSASKWADRLLSDENTSDSSSS 60  
 Db 1 MERALSHVGIIGLSPPFOLCRPPLPATRKRSRNTTCSASKWADRLLSDENTSDSSSS 60  
 Db 601 ETTVEMSVADMKEASVTKILAGVAGLISLFSQKYFLKSSSFORKDMVMSMESDVTI 660  
 Qy 601 ETTVEMSVADMKEASVTKILAGVAGLISLFSQKYFLKSSSFORKDMVMSMESDVTI 660  
 Db 661 GSVRADSEALPRMDARTAENTVKWOKIKSLAFGPDRIRLEMPEVLDGRMLKIKWDRAA 720  
 Qy 661 GSVRADSEALPRMDARTAENTVKWOKIKSLAFGPDRIRLEMPEVLDGRMLKIKWDRAA 720  
 Db 721 ETRQGLGVYDYLKUSVTSVTSADGTRALVATEBESACISDLVIRENNATDVRYTT 780  
 Qy 721 ETRQGLGVYDYLKUSVTSVTSADGTRALVATEBESACISDLVIRENNATDVRYTT 780  
 Db 781 RYEVFKSGWKITEGSVLAS 801  
 Qy 781 RYEVFKSGWKITEGSVLAS 801  
 Db 781 RYEVFKSGWKITEGSVLAS 801  
 RESULT 3  
 ID ADJ38211 standard; protein; 801 AA.  
 AC ADJ38211;  
 DT 06-MAY-2004 (first entry)  
 XX Plastid division-related Arc6 protein orthologue 4.  
 DE Prokaryotic type; plastid division; Ftn2; ARC6; Fzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target.  
 XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX WO200401003-A2.  
 XX 31-DEC-2003.  
 XX 20-JUN-2003; 2003WO-US019536.  
 PR 20-JUN-2003; 2002US-039010P.  
 PR 09-AUG-2002; 2002US-0402242P.  
 PR 20-JUN-2003; 2003US-0600070.  
 XX (UNMS ) UNTV MICHIGAN STATE.  
 PT Ostert young KW, Vitha S, Koksharova OA, Gao H;  
 XX WPI; 2004-082486/08.

PT New isolated Ftn2, ARCs and/or Fzo-like nucleic acid sequences, useful for further characterizing plastid division in plant cells, and in varying agronomic and horticultural characteristics of economically important plants.

PT Disclosure; Fig 8; 287pp; English.

CC This invention relates to novel prokaryotic type or plastid division and related genes and proteins. In particular, the invention relates to novel

Ftn2 (ARC6), ARCs and Fzo-like genes and polypeptides. The methods and compositions of the present invention are useful for further characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is that of a protein which is related to the invention.

XX SQ Sequence 801 AA;

Query Match 99.7%; Score 4051; DB 8; Length 801;  
 Best Local Similarity 99.8%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;  
 Matches 799; Conservative 0;

Oy 1 MERALSHVGIIGLSPPFOLCRPPLPATRKRSRNTTCSASKWADRLLSDENTSDSSSS 60  
 Db 61 FATATTATVLSPPSIDRPERHVPPIFYQVGLQTHFLTDGIRRAVARVKSPQFG 120  
 Qy 61 FATATTATVLSPPSIDRPERHVPPIFYQVGLQTHFLTDGIRRAVARVKSPQFG 120  
 Db 121 FSDDALISRQLQACETLSNPSRSRBYNEGCLLDBBATITDVWDKVGACLVQEG 180  
 Qy 121 FSDDALISRQLQACETLSNPSRSRBYNEGCLLDBBATITDVWDKVGACLVQEG 180  
 Db 181 GETTIVLRVGEALKERLKSKFSKQDVLVLMALAFDVSRSRDAMALDPDFITQIEFVEAL 240  
 Qy 181 GETTIVLRVGEALKERLKSKFSKQDVLVLMALAFDVSRSRDAMALDPDFITQIEFVEAL 240  
 Db 241 KLIQEGRSISLAPPRLAQIDEETEITPRYVIELLPLGGDYAARKRLNGLSGVNLWS 300  
 Qy 241 KLIQEGRSISLAPPRLAQIDEETEITPRYVIELLPLGGDYAARKRLNGLSGVNLWS 300  
 Db 301 VEGGGASALVGGSLTREKMEANFLRMTAAEQDVLFATPSNIPASESEVYEVYALVAQA 360  
 Qy 301 VEGGGASALVGGSLTREKMEANFLRMTAAEQDVLFATPSNIPASESEVYEVYALVAQA 360  
 Db 361 FIKKKPHILODAKQFOOLQOAKWMAMIPAMLYDRNWEIDFGLERGCALLIGKVE 420  
 Qy 361 CRMWGLGDSEDSQYRNPAIVFLENNSNRDDNDLPGICKLETLWLAGVPRFRDKD 480  
 Db 421 CRMWGLGDSEDSQYRNPAIVFLENNSNRDDNDLPGICKLETLWLAGVPRFRDKD 480  
 Qy 481 KERFLGDXYDDPMVLSYLVERVEVQGSPLAATAATMARIAGAHTVKASAMQALQKVFPSPYTD 540  
 Db 481 KERFLGDXYDDPMVLSYLVERVEVQGSPLAATAATMARIAGAHTVKASAMQALQKVFPSPYTD 540  
 Qy 541 RNSAEPKDVOETVFSVPVGNNGRDCGPVFTAEARVPPSENFENDYAIRGVSSSSVD 600  
 Db 541 RNSAEPKDVOETVFSVPVGNNGRDCGPVFTAEARVPPSENFENDYAIRGVSSSSVD 600  
 Qy 601 ETTVEMSVADMKEASVTKILAGVAGLISLFSQKYFLKSSSFORKDMVMSMESDVTI 660  
 Db 601 ETTVEMSVADMKEASVTKILAGVAGLISLFSQKYFLKSSSFORKDMVMSMESDVTI 660  
 Qy 661 GSVRADSEALPRMDARTAENTVKWOKIKSLAFGPDRIRLEMPEVLDGRMLKIKWDRAA 720  
 Db 661 GSVRADSEALPRMDARTAENTVKWOKIKSLAFGPDRIRLEMPEVLDGRMLKIKWDRAA 720  
 Qy 721 ETRQGLGVYDYLKUSVTSVTSADGTRALVATEBESACISDLVIRENNATDVRYTT 780  
 Db 721 ETRQGLGVYDYLKUSVTSVTSADGTRALVATEBESACISDLVIRENNATDVRYTT 780  
 Qy 781 RYEVFKSGWKITEGSVLAS 801  
 Db 781 RYEVFKSGWKITEGSVLAS 801

RESULT 4  
 ADJ38209





Qy 181 GEBIVLVRGEALKERLKPFSKPDVULMALPLDVSDAMALDPPRPTGEGFVEAL 240  
 Db 181 GEBIVLVRGEALKERLKPFSKPDVULMALPLDVSDAMALDPPRPTGEGFVEAL 240  
 Qy 241 KLIQEGGASSLADPLRAQIDETEITPRYVLELLGLPIGDYAAKRNLGSVRNIMS 300  
 Db 241 KLIQEGGASSLADPLRAQIDETEITPRYVLELLGLPIGDYAAKRNLGSVRNIMS 300  
 Qy 301 VGGCGASALVGGTREKPNNEAFL 324  
 Db 301 VGGCGASALVGGTREKPNNEAFL 324

RESULT 7  
 ADDJ38247 standard; protein; 768 AA.

ID ADDJ38247;

XX XX

XX 06-MAY-2004 (first entry)

DE Plastid division-related Arc6 orthologue protein 35.

XX KW prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target.

OS Unidentified.

PN WO200401003-A2.

PD 31-DEC-2003.

XX PF 20-JUN-2003; 2003WO-US019536.

PR 20-JUN-2002; 2002US-002242P.

PR 20-JUN-2003; 2003US-00600070.

XX (UNMS ) UNIV MICHIGAN STATE.

PI Oteroyoung KM, Vitha S, Koksharova OA, Gao H;

DR WPI; 2004-082486/08.

DR N-PSDB; ADDJ38246.

XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful for further characterizing plastid division in plant cells, and in varying agronomic and horticultural characteristics of economically important plants.

PT Disclosure; Fig 8; 28pp; English.

XX This invention relates to novel prokaryotic type or plastid division and related genes and proteins. In particular, the invention relates to novel Ftn2 (ARC6) ARCS and Fzo-like genes and polypeptides. The methods and compositions of the present invention are useful for further characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is that of a protein which is 18 amino acids long.

SQ Sequence 768 AA;

RESULT 8  
 ADDJ38274 standard; protein; 789 AA.

ID ADDJ38274;

AC AC

XX 06-MAY-2004 (first entry)

DE Plastid division-related Arc6 orthologue protein 58.

XX KW prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; Plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target.

OS Trichodesmium erythraeum.

Qy 144 RSRREYMEGLL-----DDEEATVIT-----DYPWDKYPGACVLUQ 178  
 Db 61 KORSTYDOLYLHAYDPDNLAALAAVQENRTESTKRSSTDTSLGIEITQDELVGALLQ 120  
 Qy 179 EGEGTEVLRVGEALKERLKPFSKPDVULMALPLDVSDAMALDPPRPTGEGFVEAL 240  
 Db 121 EUGYEYLVKLGRPRPYLNKNSATSSRSKNSNLADEETYESAEHPDVLTVALACLEGQ 180  
 Qy 221 -----AMALDPPDFITGVBFVEALKULIQBEGGASSLADPLRAQIDETEITPRYV 271  
 Db 181 WOOGHYENMAISLE----TGQB----LAVREG--LAFSSIQABIQDLYKPYR 225

Qy 272 LELIGLPIGDYAAKRNLGSVRNIMS-VSG-GGASALVGGTRE---KFVNNEAFLMT 327  
 Db 226 LEILALP---QEKTAAERSQGELIQLNLELDGGIDTNNDSEGINIDPLRFIQOLRHLI 283

Qy 328 AAEQVDIFPATSNIPASPESEPEYEVALLVQAFQKPKPHQDAKQFQQLQQAKUM 387

Db 284 VAHQHKUFEAOER--SSAVATLVAVALLAARGFAQDPROALTRQAROMLVRLGKQ--- 337

Qy 388 EIPAMLYDTTRNWEIDFGLERGICALLIGKVDECWRMLGUDSESDQRNPATVEFLNS 447

Db 338 -----DTHLEQSCALLGQBEATRYLESE--YE--ALAIREKS 376

Qy 448 NRUDNDLPLPGCKULETMAGSYVPERTRDKOKCFKLGDYDDPMWLSYLEREVVQGP 507

Db 377 -QSPDLPLPGCLYAEQWLOHVFPHFRDLANOQAFKLQYFANQVOYLE----- 426

Qy 508 LAAATMARBIGEHWKASAMQALQKOPVSPRYDNRNSKEPK-----DVOVTWFSV 556

Db 427 --ALPTDQTTNEWAVINPOVFOAKAGNTHFNNSKRTSASFNFHSRVNPNDLPET--- 480

Qy 557 DPVGANNVORDGSPGVFTAEAVPSENFTENDYAIRAGVSESSVDETVMSTADMUK-- 613

Db 481 -----PTK--ETSEVNPFSPPMWSSSGSTIKSEVPAERMRSRT 516

Qy 614 ---EASVTKLAAQ----- 623

Db 517 NOHLNGSAKSAHSGNOKRRRKPTPSASRERIPDORPHSKRPRRRRTFAMIEKGTRLV 576

Qy 624 --VAIGLISL-----FSQKFLKSSSFORKDWMSSMSESDVATIGSVRADSE 669

Db 577 WRVFISLVLVFWVFLRATTGWLKNLPPQPSPPDQLFLVQINQPLPFLPDRPKEST 636

Qy 670 ALPRMDARTAENIVSKQOKIKLAFGDHRIELPVIDGMLKIMWDRAETAQGLYV 729

Db 637 ECPBLTNAE-AEVITMLSTKOAALGSHENHNLQBLTGLSOMR-LIAQONKLDNR 694

Qy 730 ---DYTILKLSDSVTASDGTRALVALTBSACISDLVHENNAADVRYTTRVFW 786

Db 695 RKEDHSKIKLESVEKIGFAD--RAAEVATVKVTOYENQFKNSND--KURVYDLIR 750

Qy 787 SISGWKITEGSVL 799

Db 751 ERGKWRQTSVW 763







Db 582 AVVQAKVDE---VEQVYRGDQLETRRDGLGLVIRYOLVRENNIWKIASLVL 630  
 RESULT 12  
 ID ADJ38132 standard; protein: 631 AA.  
 XX ADJ38132;  
 DT 06-MAY-2004 (first entry)  
 DB Synechococcus ftn2 protein sequence.  
 KW prokaryotic type; plastid division; Ftn2; ARC6; Fzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target.  
 OS Synechococcus sp.  
 XX WO200401003-A2.  
 PD 31-DEC-2003.  
 PA 20-JUN-2003; 2003WO-US019336.  
 PR 20-JUN-2002; 2002US-0390140P.  
 PR 09-AUG-2002; 2002US-0402242P.  
 PR 20-JUN-2003; 2003US-00600070.  
 XX (UNMS ) UNIV MICHIGAN STATE.  
 XX Osteryoung KW, Vitha S, Koksharova OA, Gao H;  
 DR N-PSDB; ADJ38131.  
 XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful for further characterizing plastid division in plant cells, and in varying agronomic and horticultural characteristics of economically important plants.  
 XX PS Claim 1; SEQ ID NO 5; 287pp; English.  
 CC This invention relates to novel prokaryotic type or plastid division and related genes and proteins. In particular, the invention relates to novel Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and compositions of the present invention are useful for further characterising plastid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is that of a protein which is related to the invention.  
 XX SO Sequence 631 AA;

Query Match 10.9%; Score 442; DB 8; Length 631;  
 Best Local Similarity 24.5%; Pred. No. 3\_3e-30;  
 Matches 189; Conservative 118; Mismatches 262; Indels 204; Gaps 30;

QY 86 IPIIDPYQVNGAOPHFLTGIRRFAEARSKPPQFGFSDALISRRQIQACETELSPRS 145  
 3 IPIIDPYKILCVGVQASADKLAESYRDRILNQSFSHEFSBLAQARROLLEAAIBLSDEQ 62  
 QY 146 RREYNE----GILDDDEATVITDVDPDKVPGALCVIIEGGEMEVIVRVEGEALI----- 194  
 63 RDRYDRPFFGGLEATEIPSELED-WORI-GALLILRIGBYDRVSQLAEELLPDYAS 119  
 195 KERLPLKSKFKODVVLVALAPLDSVRDAMALDPDPDITQEFVEBALKUQOBSCA---- 248  
 120 AERDQFARGDIALIAALSOQSGRECRQ-----OGLYEGDQAHQHGRS 162  
 QY 249 -SSLA-----FDLRAQIDETLBSTTPRYVLELLGLPLGDDYAAKRNLNSLGVNL--W 299

Db 163 OSALADHORPPPELSRTLHQECQQLRPRILERIAQOPTAD-SDROQGULLIQAMDDQ 220  
 QY 300 SVGGGGASALVSGLTRKFMNARFLR----NTAABGDYLF---VAMPNTPAESFREVY 351  
 DB 221 GIEGPGBDGG--SGTITIONPL-MFLQOIRGYLTABEQLLSEBARRPS--PRAF---F 271  
 QY 352 VALVALVAQAPIKGKPHILQDAKQFOOLQOAKVMAMBIPAMLYDTRNWEFLDFGLERGLC 411  
 DB 272 ACVTLLARGFCOHQPSLJHRASILLHLKS----- 312  
 QY 412 ALIGKVDECRMWLGSED---SQRNPAVFEVILENSNRDDNDOLPGLCKULETWLA 467  
 DB 313 SUIGQPBEEAELL-VSQDEETLSQIRALAOGEAL---IVGLCRPTEWLA 360  
 QY 468 GVWPFPRDTKOKKFKGDDYYDDPMVLSYLERVEVYGSPLAATMARGAEGHVKSAM 527  
 DB 361 TKVFPDRDLKERTAPIQYFDPDVPDYLDAVEL----- 396  
 QY 528 QALQKVFPSPRYTKRNSAEPKDQVQETVSVDP-----VGNNVGRDQECPVIAEVPRS 580  
 DB 397 -----PSDLMPTEPLVPELVEVRSSLLAKELPPTATPG-----VAPP 432  
 QY 581 EINFETNDYAIRAGVSESSVDETTEVMSVADMKEASYKILAGVAIGLISLSFSQVKFLK 640  
 DB 433 PRRRRDRSERPARTAKGLPLFWIGLV-----VVNLGGGTGV-----WAVRS 475  
 QY 641 SASFORKDMVSSMESDVATIGSVRADSEALP-----RMDARTAENTVSKWOKIKS 691  
 DB 476 RSN-----STPPTPPVQTLPEAVPAPSPAPVTAIDLRAOPETVQNLWAKKA 524  
 QY 692 LAFGPDHIREMLEPVEUDGRMLKWTDAETA--QGLIVYDYLKLVSVDVSYVSADGR 749  
 DB 525 AALGPQTDRLATVLTGVEVLOTWQGSSQQANTOLTSQFDH--KLTVDVSVOLSDQDQR 581  
 QY 750 ALVEATELBSACSLDLYHPENNADYF---VTRIVEVFWSKSGWKITEGSVL 799  
 DB 582 AVVQAKVDE---VEQVYRGDQLETRRDGLGLVIRYOLVRENNIWKIASLVL 630

RESULT 13  
 ID ADJ38240 standard; protein: 631 AA.  
 XX ADJ38240;  
 DT 06-MAY-2004 (first entry)  
 XX Plastid division-related Arc6 orthologue protein 31.  
 DB Prokaryotic type; plastid division; Ftn2; ARG6; ARCS; Fzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target.  
 OS Synechococcus sp.  
 XX WO200401003-A2.  
 PD 31-DEC-2003.  
 PA 20-JUN-2003; 2003WO-05019336.  
 PR 20-JUN-2002; 2002US-0390140P.  
 PR 09-AUG-2002; 2002US-0402242P.  
 PR 20-JUN-2003; 2003US-00600070.  
 XX (UNMS ) UNIV MICHIGAN STATE.  
 XX Osteryoung KW, Vitha S, Koksharova OA, Gao H;  
 DR N-PSDB; ADJ38241.

PT New isolated ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful  
PT for further characterizing plastid division in plant cells, and in  
PT varying agronomic and horticultural characteristics of economically  
PT important plants.

PS Disclosure; Fig 8; 287pp; English.

XX This invention relates to novel prokaryotic type or plastid division and  
CC related genes and proteins. In particular, the invention relates to novel  
CC ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and  
CC compositions of the present invention are useful for further  
CC characterising plastid division in plant cells, in order to vary  
CC agronomic and horticultural characteristics of economically important  
CC plants, such as crop, ornamental and woody plants. They can also be used  
CC as herbicide targets. The present sequence is that of a protein which is  
CC related to the invention.

SQ Sequence 631 AA;

XX

**RESULT 14**  
 ADD38253 ID ADJ38253 standard; protein; 819 AA.  
 XX  
 AC ADD38253;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DR Plastid division-related Arc6 orthologue protein 40.  
 XX  
 KW prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;  
 KW agronomic; horticultural; crop plant; ornamental plant; woody plant;  
 KW herbicide target.  
 XX  
 OS *Arabidopsis thaliana*.

Matches	Best Local Similarity	Local Similarity	Conservative	Pred.	No.	Date	3-jc-30	Indels	Gaps
QY	86	IPIDPYQVQGAQOHTFLTDGTRRAFARVSKEPQFGFSDDALISRQILQACETSNPRS	145						
Db	3	IPIDPYQVQGAQOHTFLTDGTRRAFARVSKEPQFGFSDDALISRQILQACETSNPRS	145	IPIDPYQVQGAQOHTFLTDGTRRAFARVSKEPQFGFSDDALISRQILQACETSNPRS	145				
QY	146	RREYNE-----GLIDDEEATVITDWPWDKVFGALCVLQEGGETEVTLVRGEALL-----	194	RREYNE-----GLIDDEEATVITDWPWDKVFGALCVLQEGGETEVTLVRGEALL-----	194				
Db	63	RDRYDRFFQGGLEAIEPSLEED--WORI-GALLILEGBYDRVSQLAELLPDYDAS	119	RDRYDRFFQGGLEAIEPSLEED--WORI-GALLILEGBYDRVSQLAELLPDYDAS	119				
QY	195	KERLKSKFKQDVVUWMAFLDVSRSRDAALDPDPDITGVFVERBALKLQOREGA-----	248	KERLKSKFKQDVVUWMAFLDVSRSRDAALDPDPDITGVFVERBALKLQOREGA-----	248				
Db	120	AEVRDQFARGDIALIAALSOOSLGRBCRQ-----	162	AEVRDQFARGDIALIAALSOOSLGRBCRQ-----	162				
QY	249	-SSLA-----PDLRAQIDSTLEETPTRYVBLGLPIGDYAAKRLNGLSGVRNIL---W	299	-SSLA-----PDLRAQIDSTLEETPTRYVBLGLPIGDYAAKRLNGLSGVRNIL---W	299				
Db	163	QSLALDHQRFPPELSRTLHQBHQQPRYPTLERLAPLTAD--SDRQQGQLLQAMLDRO	220	QSLALDHQRFPPELSRTLHQBHQQPRYPTLERLAPLTAD--SDRQQGQLLQAMLDRO	220				
QY	300	SVGGGGASALVGGETTREKEMNEAFR-----MTAALQVDIF--WATPSNTPAESFEVYE	351	SVGGGGASALVGGETTREKEMNEAFR-----MTAALQVDIF--WATPSNTPAESFEVYE	351				
Db	221	GIEGPDDG--SGLTLDNFL--WFLQIQRGYLTLAQOLLFSESEARRPS--PAASF--F	271	GIEGPDDG--SGLTLDNFL--WFLQIQRGYLTLAQOLLFSESEARRPS--PAASF--F	271				
QY	352	VALALVAQAFIGKKEPHILQDADKQFQLQQAKVAMEIPAMLYDTRANNWFELDGCLERGLC	411	VALALVAQAFIGKKEPHILQDADKQFQLQQAKVAMEIPAMLYDTRANNWFELDGCLERGLC	411				
Db	272	ACTYLTAQFCDCDHSPLHRSALLHLK-----	312	ACTYLTAQFCDCDHSPLHRSALLHLK-----	312				
QY	412	ALLIGKVDDECMMWGMILDSED---SQRYNIAIVEVTLENSNRDDIDLPGCLKLETWLA	467	ALLIGKVDDECMMWGMILDSED---SQRYNIAIVEVTLENSNRDDIDLPGCLKLETWLA	467				
Db	313	SLLGQPEBAAELL-VQSQDEETUSQIRALAQGEL-----	360	SLLGQPEBAAELL-VQSQDEETUSQIRALAQGEL-----	360				
QY	468	GUVPFRPRFTDKKCKPKLGDXYYDDPWNLYSLVERVEYQGSPLAAAATMARIQAEHYKASAM	527	GUVPFRPRFTDKKCKPKLGDXYYDDPWNLYSLVERVEYQGSPLAAAATMARIQAEHYKASAM	527				
Db	361	TKVFPDFRDLKERTAPLQSYFDDPQVQYDAIVL-----	396	TKVFPDFRDLKERTAPLQSYFDDPQVQYDAIVL-----	396				
QY	528	QALQKVPSRTYDANSAEPKDVQEVTSVSDP-----VGNNVGDRGEPEVIAEVRRPS	580	QALQKVPSRTYDANSAEPKDVQEVTSVSDP-----VGNNVGDRGEPEVIAEVRRPS	580				
Db	397	-----PSDLMPFLPVPVPLEVURSSLAKELPPATPG-----VAPP	432	-----PSDLMPFLPVPVPLEVURSSLAKELPPATPG-----VAPP	432				
QY	581	ENFBTDYAIRAGVSSESSDETTEMSVADMKEAVKSVILLAAGVAGLISLFSQXFLKS	640	ENFBTDYAIRAGVSSESSDETTEMSVADMKEAVKSVILLAAGVAGLISLFSQXFLKS	640				
Db	433	PRRRDRDRSERPARTAKRPLPMWIGGV-----VVVIGGTCV-----WAWS	475	PRRRDRDRSERPARTAKRPLPMWIGGV-----VVVIGGTCV-----WAWS	475				
QY	641	SSSFQKDMVSSMSESDVATIGSVVRADDSEALP-----RMDARTAENTVSKOKIKS	691	SSSFQKDMVSSMSESDVATIGSVVRADDSEALP-----RMDARTAENTVSKOKIKS	691				
Db	476	RSN-----STPPPPPVVQTLPEAVPAPSAPVTVLQDQAETVLQNLAAKA	524	RSN-----STPPPPPVVQTLPEAVPAPSAPVTVLQDQAETVLQNLAAKA	524				
QY	692	LARGPDHRLEMPLPVTFDGRMLKIKWMDRASTA--OGLGVVDTULKLSVSUTVADGTR	749	LARGPDHRLEMPLPVTFDGRMLKIKWMDRASTA--OGLGVVDTULKLSVSUTVADGTR	749				
Db	525	ARLGPOYDRDRDLATVLTGEVLIQTOGCGFSQOANTTSQFDH--KUTVDSVLSQGDQR	581	ARLGPOYDRDRDLATVLTGEVLIQTOGCGFSQOANTTSQFDH--KUTVDSVLSQGDQR	581				
QY	750	ALVATELTERSACSLDVHNENNADVRT--YTTRYEVFVFSKSGKMKTECSV	799	ALVATELTERSACSLDVHNENNADVRT--YTTRYEVFVFSKSGKMKTECSV	799				

Db 240 LARAQSLIKSKVTLGKLA-LITQIESLEGGLAPPCTIDLGLIPRTPENAERRGAIAL 297  
 Qy 295 RMLIWSVGGGASALVALGVLTRK-----FMNRAFLRMTAEOVDLF----VATP 339  
 Db 298 REILRQ-----GLSVERASCQIQDWPWPCFLSQASRLIRATEVLDLPPWDLAIKR 346  
 XX PT  
 Qy 340 SNPAESFE-----VYEVALVALVAQAFIGKKPHLQDAKQFOQQAKVAMEIP 390  
 :  
 347 NKKSLESHNRQVIVIDENCFCYMLIGHAVGFSG-----KONETINKAKTICBCL 396  
 Db 391 AMLYDTTRNNWEIDFGLERGLCALLIGKUDCECRMWGL---DSEDSQRNPNAIVEVLEN 446  
 A-----SEGDLKFEAFCSFLIKQGSEARALEKLKOLENSDSAVRNS----ILGK 444  
 397 :  
 Db 447 SNRDDNDLPGLKQKLETLWLAGVVFPRFRDTK-----DKKPKLGDDYDDPMVL 494  
 :  
 445 ESRSTS-----ATPSLBAWLMESVLANFPDTGCSPLANFRAEKYPENKGSPSIM 499  
 Qy 495 SY-----LERVEVQOSPLAAATMARIAGAEHKASAMQAQKVFSRVTDRNSAEPKD 548  
 ::  
 500 NHKTNRQPLSTTQFVNSS-----OHL-----YTAVEQLPTD 531  
 Db 549 VOTEVFSVPVGANNVRGRDGEPEGVIAEAVRPSENFTNDYAIRAGVSSESV-DETVEMS 607  
 Qy :  
 Db 532 LOSPVVSAK--NNDE-----TSASMSVQVKRN----LGVHNKWDW--LS 571  
 Qy 608 VADMKEASVKILAGVAIGLISLFSQRY----- 636  
 Db 572 OSSLIGRVSVVVAL-----LGCTVFFSUKLGSIGRSRQLQSMPIISVAPRHESSEDFLWKT 626  
 Qy 637 --PLKSSSFQPKDMVSS-----MESUVATGGSVRADSEALP 673  
 :  
 Db 627 SGPNRKNLDSVNRNGIVNIKILIDMLKHCGHRPDALYKLSSQGOSATSLSHASELHKR 686  
 Qy 674 -MDARTENIVSKWQKISLACPDPHRLEMPLVFDGMLKWTWDRATAETQGLVW-DY 731  
 :  
 Db 687 PHDTEEARBLVQWEMNTKARALGQPHQVSYLSBVLDLSDMVLQW-QTIAQTAARKSCYWRF 745  
 Qy 732 TIKLKLSDVSVTSSAD--GTRALVATELEBESACLSDLVPHENNATDVYTTRYEVFSK 788  
 :  
 Db 746 VILHLLEVLOQAHFEDGIGAGEAABEIALEELAEALVDESQPK-NAKYVSTYKURYLKKQE 804  
 Qy 789 SG-WKLTGEGSV 798  
 Db 805 DGLWFKCQSDI 815  
 RESULT 15  
 ADJ38251  
 ID ADJ38251 standard; protein; 819 AA.  
 XQ  
 AC ADJ38251;  
 XQ  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Plastid division-related Arc6 orthologue protein 39.  
 XX  
 KW prokaryotic type; plastid division; Ftn2; ARC6; Fzo; plant cell;  
 KW agronomic; horticultural; crop plant; ornamental plant; woody plant;  
 KW herbicide target.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200401003-A2.  
 XQ  
 PD 31-DEC-2003.  
 XX  
 PR 20-JUN-2003; 2003WO-US019536.  
 XX  
 PR 20-JUN-2002; 2002US-0390140P.  
 PR 09-AUG-2002; 2002US-0402242P.  
 PR 20-JUN-2003; 2003US-00600070.  
 XX  
 PA (UNMS ) UNIV MICHIGAN STATE.  
 XX PT Osterryoung KW, Vitha S, Koksharova OA, Gao H;  
 XX DR WPI; 2004-082486/08.  
 XX PT New isolated Ftn2, ARC6 and/or Fzo-like nucleic acid sequences, useful  
 PT for further characterizing plastid division in plant cells, and in  
 PT varying agronomic and horticultural characteristics of economically  
 PT important plants.  
 XX RS Disclosure; Fig 8; 287pp; English.  
 XX Sequence 819 AA:  
 Query Match 10.6%; Score 431; DB 8; Length 819;  
 Best Local Similarity 23.5%; Pred. No. 5e-29;  
 Matches 20; Conservative 116; Mismatches 311; Indels 224; Gaps 32;  
 Matches 20; Conservative 116; Mismatches 311; Indels 224; Gaps 32;  
 Db 04 56 SSSSFATTATTATLVSLLPSPSTRPERHRVPIFQVQLGAQTHFELDGIRAFEARVSK 115  
 Qy 81 SRTSSLAST----- 124  
 Db 116 PROGFSDDALISRROIIQAOACTLSNPRSRREYNGSLLDDEBEATVTDVFWDKVSGALC 175  
 Qy 125 DABEGYTMEEAAARQDILMDVYDRL---LFFSBSYAGNLKEKIAPKPLRITIWAWLGALC 181  
 Db 176 VIOEGGETTEIVRVEAGLKERLKPFSKQDVYLMALFLVSDRDAALDPDFITGYEF 235  
 Qy : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| :  
 182 LIQEVSGEKLVLIDIGRALRNLDSKPYIHDFITLMSMCALECAIAKAFAEVNKVS--QGPEA 239  
 Db 236 VEBALKLILQEE-GASSLAPDRAQIBETLEETITPRVLELIGLIPGLGDYAKRNLNSGV 294  
 Qy :  
 Db 240 LARAQSLIKSKVTLGKLA-LITQIESLEGGLAPPCTIDLGLIPRTPENAERRGAIAL 297  
 Qy 295 RMLIWSVGGGASALVALGVLTRK-----FMNRAFLRMTAEOVDLF----VATP 339  
 :  
 Db 340 SNPAESFE-----VYEVALVALVAQAFIGKKPHLQDAKQFOQQAKVAMEIP 390  
 :  
 347 NKKSLESHNRQVIVIDENCFCYMLIGHAVGFSG-----KONETINKAKTICBCL 396  
 Db 391 AMLYDTTRNNWEIDFGLERGLCALLIGKUDCECRMWGL---DSEDSQRNPNAIVEVLEN 446  
 A-----SEGDLKFEAFCSFLIKQGSEARALEKLKQOLNSDSAVRNS----ILGK 444  
 397 :  
 Db 447 SNRDDNDLPGLKQKLETLWLAGVVFPRFRDTK-----DKKPKLGDDYDDPMVL 494  
 :  
 445 ESRSTS-----ATPSLBAWLMESVLANFPDTGCSPLANFRAEKYPENKGSPSIM 499  
 Qy 549 VOTEVFSVPVGANNVRGRDGEPEGVIAEAVRPSENFTNDYAIRAGVSSESV-DETVEMS 607  
 :  
 Db 532 LOSPVVSAK--NNDE-----TSASMSVQVKRN----LGVHNKWDW--LS 571  
 Qy 608 VADMKEASVKILAGVAIGLISLFSQRY----- 636  
 Db 572 OSSLIGRVSVVVAL-----LGCTVFFSUKLGSIGRSRQLQSMPIISVAPRHESSEDFLWKT 626

us-10-600-070-2.rag

QY  
 Db  
 637 ---FLKSSSSFORKOMVSS-----  
 627 SGNFRKNLDSVNRRNGIVGNKVLIDMLKNCHGCEHDALYLIKSGSOSATSLSHSASBLHKR 686  
 QY  
 674 -MDARTAENIVSKKKIKSIAFGPDRPIELPEVDGMRKLIWTBRAETRQLGLVY-DY 731  
 Db  
 687 PMDTEEBELVRQMDENVKABALGPHTQVSLESEUDESMUVQW-QTLaQTAEEAKSCWRF 745  
 QY  
 732 TLLKLSVSVTSVSAAD--GTRALVATELBESACSLDVHPENNATDVTYTRYEVWSK 788  
 746 VLLHLLEVLOMHIFBDGIAGERAEEIRALLEPARELVEDESOPK-NAKYSTYKIRYLKQOE 804  
 QY  
 789 SG-WKITECSV 798  
 Db  
 805 DGLWKFCOSDI 815

Search completed: June 10, 2005, 01:38:50  
Job time : 101 secs

THIS PAGE BLANK (USPTO)



Db 1913 RAGYRPSALSGSATGVILGSMGSYDGLAHITVDIKELDGYRGIGSAASISG-----R 1964  
 REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 303-231-1098  
 TELEX: 1965 VAKVAMMAMETPAMYDTRNNWEITFGLCALLIGKVDECRMWGLGJLSDQSYRNPAI- 439  
 INFORMATION FOR SEQ ID NO: 23:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2089 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: Single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-418-893D-23

Db 2064 GDRVLGVIRGSANQDGSRQSGLTAPNGPAQQRTVRQALSSCGISPEDIDAVENHTGTSL 2123  
 QY 505 GSPLLAAATMARGAE-----HVKASAMOA--LOKVPSPRTDRNDAE 545  
 Db 2124 GDTIEAGLAEVTCGERSPERPYLYGSSRSNLSQVAKVAVKML-----SMQ 2175  
 QY 546 PKVQVETPSVDPVGNNVRGDPGPGVFTIAAVTPSSENTEINDAIRAGVSSESSDTEVE 605  
 Db 2176 HEVLPKPTMAEAO--SPHGWEGS--GLSLQEAP--WRRNGSVRAGVSSFCISGTMH 2230  
 QY 606 MSVADMLKEA--SVKILLAAGVAGLISLFSQKYLKLSSSFORKDMVSSME---SDVAT 659  
 Db 2231 IIDEAAPAEARRPVEEAAPALLPLVLSGRDAVNQAGRMWKMLBEEHGEVGWSDVR 2290  
 QY 660 IGSVRADSEALPMDARTAENTVSKWOKIKSAFG--PDHRIMPEVTDGRMKINTDR 718  
 Db 2291 TAKLHRTHFESRASVLSAAGAV--BGRRAASSGRDAVSGTAKRGKGKAVALPTGQ 2347  
 QY 719 AAETAQLG--LVVDTYLLKLSVSVTVSAD----GTRALVEATL--BESACU--SDLVH 767  
 Db 2348 GSCORLGKMRKVLYVPPVRAAFDCEVALEDAYLDGRGLREVFAAGSBEGALLERTEVTO 2407  
 QY 768 PENNADVRYTYTRYEVWKSQSWK 792  
 Db 2408 PGIAPEALVALYRQ----WESWQIK 2427

RESULT 2

US-08-418-893D-23  
 Sequence 23, Application US/08418893D  
 Patent No 559220  
 GENERAL INFORMATION:  
 APPLICANT: ROESSLER, PAUL G  
 APPLICANT: OHLROGGE, JOHN B  
 TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A  
 NUMBER OF SEQUENCES: 25  
 TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTELLA CRYPTICA  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY  
 STREET: 1617 Cole Blvd.  
 CITY: Golden  
 STATE: CO  
 COUNTRY: USA  
 ZIP: 80401-3193  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/418,893D  
 FILING DATE: April 7, 1995  
 CLASIFICATION: 800  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/104,938  
 FILING DATE: September 14, 1993  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: O'CONNOR, EDNA  
 REGISTRATION NUMBER: 29,252

Db 530 -----LQKVPSPRTDRNDAEPEQDVET---- 552  
 QY 645 TAGIPGINSNENTEVAYLDTYPFHVERISPDVYFTLGDNTIDVEVTQGALLATFGG 704  
 Db 553 ---VFSID-PVGNNVRGDB---PGVTAEVRES----ENFTNDVYARAGVS 597  
 Db 705 ETRIFGMDEPLGRLSLGATVLMPTFDISBELRTDVTGKVRYQDNGATVEAG--- 760  
 QY 598 SYDFTTVMVSADMLKEAASKVILLAAGVAGLISLFSQKYLKLSSSFORKDMVSSME---SDVAT 659  
 Db 761 ---OPYVVEAMKMI--MPIKATESG----KITHNISAGSVTISAGOLIASLBDK 806  
 QY 657 VATIGSV---RADSBALPRMDARTA-ENVSKWOKIKSAFGPDHRIMPEVTDGR 710  
 Db 807 PERVKKIEFTESKLIDNESKVNDEPKAVMVNL----GFNLD----PEAVAQ 852  
 QY 711 MLKIKWDRDRAETAQOLGLVYD-YTLKLKLSVDSVTSADGTRALVEATLBBACISDLVHP 769

**RESULT 3**  
 US-08-418-893D-24  
 ; Sequence 24, Application US/08418893D  
 ; Patent No. 555220  
 ;  
 APPLICANT: OHLROGGE, JOHN B  
 TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A  
 TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTELLA CRYPTICA  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY  
 STREET: 1617 Cole Blvd.  
 CITY: Golden  
 STATE: CO  
 COUNTRY: USA  
 ZIP: 80401-3393  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/418, 893D  
 FILING DATE: April 7, 1995  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: O'CONNOR, EDNA  
 PRIOR APPLICATION DATA:  
 REGISTRATION NUMBER: 29, 252  
 APPLICATION NUMBER: US 08/104, 938  
 FILING DATE: September 14, 1993  
 CLASSIFICATION: 800  
 REGISTRATION NUMBER: MRL/NREL IR# 92-48CON  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 303-231-1000  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 24:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2089 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULAR TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 US-08-418-893D-24

Query Match Score 143; DB 1; Length 2089;  
 Best Local Similarity 20.0%; Prod. No. 0.0009;  
 Matches 169; Conservative 113; Mismatches 289; Indels 272; Gaps 44;

**RESULT 4**  
 US-09-057-969-2  
 ; Sequence 2, Application US/09057969  
 ; Patent No. 6013451  
 ;  
 GENERAL INFORMATION:  
 APPLICANT: VICTOR THI WONG  
 APPLICANT: WONG, VICTOR THI WONG  
 APPLICANT: PHANG, SENG MENG  
 APPLICANT: TAN, TIEN CHIE  
 TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA  
 TITLE OF INVENTION: POLYMERASE I (KLENOW) CLONES INCLUDING THOSE REDUCED TO  
 TITLE OF INVENTION: 3'-TO-5' EXONUCLEASE ACTIVITY  
 NUMBER OF SEQUENCES: 17  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
 STREET: 745 FIFTH AVENUE  
 CITY: NEW YORK  
 STATE: NEW YORK  
 COUNTY: USA  
 ZIP: 10151  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

Qy 227 PDRFTGYVERVEEALKLQOBEGASSLAAPDRAQIDETLEEITPRVYLELGLPLGGDDYAK 286  
 Db 358 RDCTQRVR---OKIFREGPSIVP---KETPHM-----ELAQ 391  
 Qy 287 RLNGLSGVNTILMSVGSGGASAALVGGLREKEFNEAFURMTAAEVDLFLFA--TPSNIP 343  
 Db 392 RL----TQNIGYQ-GAGTVEVLNAAKNEPKDQADKQFCQLQQAKUMAMEIPA-- 391  
 Qy 344 AES----PEVYEVIALQAQFICKKKPHILQADKQFCQLQQAKUMAMEIPA-- 391  
 Db 442 ATOLQVAMGIPLENIPDRLRKYRDRDAYGTDP--IDFLQERTREL-DSHVIAARTAPN 498  
 Qy 392 -----MLYDPTRNMM-BIDFGLERGICALLIGKVDECRMNLGLSBDQY-- 434  
 Db 499 DEGFKPTGSIERIKIKFQSTPNVNGYFSVCGANGGI----HERADSQRGH 542  
 Qy 435 --RNP-----AIVERTLENSNRD-NDPLGCLKIET-----WLAGUVFP 472  
 Db 543 LFAKGPNRREOARKALVALKEMEVGRDTRNSVELVLUKLIETEAFKNTIDTSNLGII-- 600  
 Qy 473 RFRDTKDKFCKLGLDYDPMVMSYLERVWVQSSPLAAATMARIAGAHVKASAMOA-- 529  
 Db 601 ----KESVKV-----EMPSHL---VVVGAVERKAHEVKVATEVKESFKRSQVS 644  
 Qy 530 -----LQKVPPSRVTDRNSAEPKVQET---- 552  
 Db 645 TAGIPGINSFNFIEVAVIDTLYKPVHVERISPDVYRFTLDGNTIDVENVOTPAEGALLATPGG 704  
 Qy 553 ---VFSVD-PVGNVNGDGE---PVVTFAEVRRPS----ENFTENDYAIRAGYSES 597  
 Db 705 EHTRIFGMDEPLGLRLSISLGATVLMPTIDFDPSERLTDTVGKVVRVYLDQNGATVEAG--- 760  
 Qy 598 SVDBTIVMSVADMILKEASVSKVLAAGVAGLSSLPSQXPKLSSSFORKDMVSSME SD 656  
 Db 761 --OPYVNETEAMONI-MPKATEG-----KITHNLRLAGDILASBLEKD 806  
 Qy 657 VATIGSV---RADDEALPRMDARTA-ENTVSKWQKIKSLAFLGPDIRIEMLPBEVLDGR 710  
 Db 807 PSRVKKISTFSKGLDIMSKVDLRQPQKAWMVL-----GFNLID----PEAVHQ 852  
 Qy 711 MLKJWTDRAETQOLGYTD-YTLKLSVDSVTSVADGTRALVETABESACLSLWHE 769  
 Db 853 AIDSATSSAADDLVQVLDDEFYRVERSESQDFGV-IADDVVRLTKANTE---TLDVWISE 907  
 Qy 770 NNA 772  
 Db 908 NLA 910

Qy 227 PDRFTGYVERVEEALKLQOBEGASSLAAPDRAQIDETLEEITPRVYLELGLPLGGDDYAK 286  
 Db 358 RDCTQRVR---OKIFREGPSIVP---KETPHM-----ELAQ 391  
 Qy 287 RLNGLSGVNTILMSVGSGGASAALVGGLREKEFNEAFURMTAAEVDLFLFA--TPSNIP 343  
 Db 392 RL----TQNIGYQ-GAGTVEVLNAAKNEPKDQADKQFCQLQQAKUMAMEIPA-- 391  
 Qy 344 AES----PEVYEVIALQAQFICKKKPHILQADKQFCQLQQAKUMAMEIPA-- 391  
 Db 442 ATOLQVAMGIPLENIPDRLRKYRDRDAYGTDP--IDFLQERTREL-DSHVIAARTAPN 498  
 Qy 392 -----MLYDPTRNMM-BIDFGLERGICALLIGKVDECRMNLGLSBDQY-- 434  
 Db 499 DEGFKPTGSIERIKIKFQSTPNVNGYFSVCGANGGI----HERADSQRGH 542  
 Qy 435 --RNP-----AIVERTLENSNRD-NDPLGCLKIET-----WLAGUVFP 472  
 Db 543 LFAKGPNRREOARKALVALKEMEVGRDTRNSVELVLUKLIETEAFKNTIDTSNLGII-- 600  
 Qy 473 RFRDTKDKFCKLGLDYDPMVMSYLERVWVQSSPLAAATMARIAGAHVKASAMOA-- 529  
 Db 601 ----KESVKV-----EMPSHL---VVVGAVERKAHEVKVATEVKESFKRSQVS 644  
 Qy 530 -----LQKVPPSRVTDRNSAEPKVQET---- 552  
 Db 645 TAGIPGINSFNFIEVAVIDTLYKPVHVERISPDVYRFTLDGNTIDVENVOTPAEGALLATPGG 704  
 Qy 553 ---VFSVD-PVGNVNGDGE---PVVTFAEVRRPS----ENFTENDYAIRAGYSES 597  
 Db 705 EHTRIFGMDEPLGLRLSISLGATVLMPTIDFDPSERLTDTVGKVVRVYLDQNGATVEAG--- 760  
 Qy 598 SVDBTIVMSVADMILKEASVSKVLAAGVAGLSSLPSQXPKLSSSFORKDMVSSME SD 656  
 Db 761 --OPYVNETEAMONI-MPKATEG-----KITHNLRLAGDILASBLEKD 806  
 Qy 657 VATIGSV---RADDEALPRMDARTA-ENTVSKWQKIKSLAFLGPDIRIEMLPBEVLDGR 710  
 Db 807 PSRVKKISTFSKGLDIMSKVDLRQPQKAWMVL-----GFNLID----PEAVHQ 852  
 Qy 711 MLKJWTDRAETQOLGYTD-YTLKLSVDSVTSVADGTRALVETABESACLSLWHE 769  
 Db 853 AIDSATSSAADDLVQVLDDEFYRVERSESQDFGV-IADDVVRLTKANTE---TLDVWISE 907  
 Qy 770 NNA 772  
 Db 908 NLA 910





Db 662 LAHIAE--DDNLLIEAFRGLDIHTKAMDIFVSEEDVTANRROQAKAVNFGIVYGISD 718  
 Qy 635 KYFLKSSSFQRKDMMSSMESDVATIGSVRADDSSEALPRMDARTAENIVSKMOKIKSLAF 694  
 Db 719 -YGLAQMINITRKEAABFERYASFPFGK-.....-QYMDNIVOE-AKQKGVT 764  
 Qy 695 GDPHRIMPLEPEVLGDRM-LKIWTDRAA-ETAOGLGLVYDYLKLSD-SVTVSADGTRA- 750  
 Db 765 TILHRRLYLPDTSRNSRVSAPARTAMNTPIQGSAAD-TIKKAMIDLSVRLEERLQAR 823  
 Qy 751 -----LVEATLERSACISDLVPE 769  
 Db 824 LLIQVHDELILBPKEEIERLCRLV-PB 850

RESULT 7  
 US-08-436-664-32  
 ; Sequence 32, Application US/08436664  
 ; Patent No. 587428  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RIGGS, MICHAEL G.  
 ; APPLICANT: SIVARAM, MATHOOR  
 ; APPLICANT: TUDOR, STARLA D.  
 ; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS  
 ; NUMBER OF SEQUENCES: 34  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Gen-Probe Incorporated  
 ; STREET: 9880 Campus Point Drive  
 ; CITY: San Diego  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 92121  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FABESBQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/436,664  
 FILING DATE: 08-MAY-1995  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/222,612  
 FILING DATE: 16-SEP-1994  
 APPLICATION NUMBER: 08/307,410  
 FILING DATE: 16-SEP-1994  
 APPLICATION NUMBER: 08/222,612  
 FILING DATE: 16-SEP-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fisher, Carole A.  
 REFERENCE/DOCKET NUMBER: GP94003.GP2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619-535-2807  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 32:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 876 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: Single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: internal  
 ORIGINAL SOURCE:  
 US-08-436-664-32

Query Match 3.2%; Score 128.5; DB 2; Length 876;  
 Best Local Similarity 21.0%; Pred. No. 0.0053;

---

Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;  
 Qy 99 HFLTDGRRRAFEARVSKEPPQFGPSDDALISRSRQLQACETISNPRSRREY-NEGILDDE 157  
 Db 107 HVEDDITGTTAAKAER- -GPFAKVISGDRDLT---LASPQTVTEITKKGTDIE 158  
 Qy 158 EAIVTVY------PFDKPG-----ALCVIQLQEGSTEIVRVG 190  
 Db 159 SYTPETVVEKYGLTPEQIVDVKLGIMGDKSNDIPGVPGIGEKTAVKLUKQFGIVENVLA 218  
 Qy 191 EALKERLKPSKQ- -QVTLWALAFLDVSDRA--MALDPDFITGYEPFEALKLOE 245  
 Db 219 DEIKGEKLUENLRQYRDLAL-LSKOLAAICRDAPVELTB--DIVYKGEDRKVVALFOE 275  
 Qy 246 EGASSLAPDLRQIDE-----TLEETPTVYLELGL--PLGDDYAAKRLNG 290  
 Db 276 LQPSFLDKMAVQTDECEKPLAGMDFAIADSVDTEMALDKAVALVWVGDVYHAPIVG 335  
 Qy 291 -----LSGVNLWLSVGGGGASALVGLGKTFKMFNEAFLRMTAEE---- 330  
 Db 336 ALLANERGRFFLRPETALADPRAK-----LGDETTRKTMDSKRAVALWKKG 385  
 Qy 331 -----QVDFIVATPSNPAES-----PEVYEVALLVAQAFIGK-----KPHLQ 370  
 Db 386 BLRGVVFDLILAYLDPQAQAGDVVAVAKHOYEAVRSDSA VYKGAKRTVDPDETLAE 445  
 Qy 371 DADKQFQOLQQAKMAMEI PAMYDTRANNWE-IDFGLERGLCALLIG----KVDCRM 423  
 Db 446 HILAR-----KALAIWALEPFLDLELRNEQDLTTELEQPLAGIANMEFTGKVVKTRL 500  
 Qy 424 -WIGLD-SEBDSYRNPAIVEFLENSPPDDNDLPGLKCLBETWLAGVVFPRFRDCK 481  
 Db 501 BOMGAELTEQLO-----AVERRTYBLAQEFNFNISP---KOJGT----VLFDKLQOLVULK 549  
 Qy 482 FKLGDYVDPMTSIL-ERVEVQGSPLAAATMARIGAHTVKASAMQALQKUVFESRYTD 540  
 Db 550 TKNG-YSTSADVLEKAPHHEVE--HILHYROLKGQUSTYLE-GLUKVWHPVTGKWT 605  
 Qy 541 RNSAEPKDVQET--VFSVDPVNNGVGRDGEPEGPVFIARARVE-----NFETNDYIA 590  
 Db 606 ENQA----LTQGRLSSVPEMLQNPPIRLEGSKRQAFVSPDWLIPADYDYSQELRV 661  
 Qy 591 RAGVSESSVDETFE-----MSVADMKE--ASVKAIIAGVAGLISLFSQ 634  
 Db 662 LAHIAE--DDNLLIEAFRGLDIHTKAMDIFVSEEDVTANRROQAKAVNFGIVYGISD 718  
 Qy 635 KYFLKSSSFQRKDMMSSMESDVATIGSVRADDSSEALPRMDARTAENIVSKMOKIKSLAF 694  
 Db 719 -YGLAQMINITRKEAABFERYASFPFGK-.....-QYMDNIVOE-AKQKGVT 764  
 Qy 695 GDPHRIMPLEPEVLGDRM-LKIWTDRAA-ETAOGLGLVYDYLKLSD-SVTVSADGTRA- 750  
 Db 765 TILHRRLYLPDTSRNSRVSAPARTAMNTPIQGSAAD-TIKKAMIDLSVRLEERLQAR 823  
 Qy 751 -----LVEATLERSACISDLVPE 769  
 Db 824 LLIQVHDELILBPKEEIERLCRLV-PB 850

RESULT 8  
 US-08-436-664-34  
 ; Sequence 34, Application US/08436664  
 ; Patent No. 587428  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RIGGS, MICHAEL G.  
 ; APPLICANT: SIVARAM, MATHOOR  
 ; APPLICANT: TUDOR, STARLA D.  
 ; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS  
 ; NUMBER OF SEQUENCES: 34  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Gen-Probe Incorporated  
 ; STREET: 9880 Campus Point Drive



LENGTH: 876 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

US-09-135-642-20

Query Match

Best Local Similarity 3.2%; Score 128.5; DB 3; Length 876;

Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

Qy 99 HFLTDGIRRRAFERARVKPPQPGFSPDALISRQLQACETLSNPRREY-NEGLIDDE 157

Db 107 RYEDDITGMARAERB--GFPAVKVSGRDLTQ---LASPQTVETTRKGIDIE 158

Qy 158 EATVITDV-----PMWKVPG-----ALCVIQEGGETEIVURVG 190

Db 159 SYTPETVEKVLTPEQIVDPLKGMDKSDNPNGVPGIGEKTAVKLJKQFGTENVIASI 218

Qy 191 EALKERKUPKSFRQ--DVTLMVALAFLDVSRLA--MALDPDFITGYEPVERALKIQLQE 245

Db 219 DEIKGEKULKENLRQYRDIAL-LSKQLAICRDAPVELTD--DIVKGEBDKVALFQE 275

Qy 246 EGASSLAPPDRQIDE-----TLEBITPRVYLELGLI--PLGDDYAAKRUNG- 290

Db 276 LQPOSFLDRMAYQVOTDEGEEKPLAGMDFAATADSVIDEMLAQDAALVVEVGDNYHAPIVGI 345

Qy 291 -----LSGVNLWMSVGGGASALVGLTREKFMEAVPLRTMAE--- 330

Db 336 ALANERGRFFELPBTALADPKELAW-----LGDETKCTMFDISKRAVALWKKG 385

Qy 331 -----QDLFVATPSNTPAES-----FEVVEVALAUVAQAFIGK-----KPHILQ 370

Db 386 ELRGVVVFDDLLAAYLDPQAQGDAVAVAKMQYEAVSDEAVYGKAKRITVPEPTIAE 445

Qy 371 DAHKOFQOLQOKYAMAMEIPAMUYDTRNWNE-IDFGJERGLCALLIG-----KVDCRM 423

Db 446 HILR-----KANAIWAERPLMDELRNNEQDRILTELEQPLAGILANNMBFTGVKVDTKRL 500

Qy 424 -WGLD-SEDSQYRNPAIVFVLENSNRDDNDLPGLKLUFTWLAGVUPPFRTDKK 481

Db 501 EQNGABELTEQLO-----AVERTRYELAQEFBNINSP---KQCLG----VLFDKQLQPVLIK 549

Qy 482 FKLGDDYDDPMWISYL-ERVEVVOGSPLAAATMARTGAEHVKAASAMQALQVFPSPRIID 540

Db 550 TKIG-YTSISADVLEKLPHETIE--HILHYRQOLQKOSTYB-GLIKVWVHTGKHTM 605

Qy 541 RNNSAEPKVOET--VFSUDPVGNVGRDGEPGVFTAEARVRPSB-----NFETNDYAI 590

Db 606 FNQK-----LTOGRLLSSVEPNLQNPILKEBEKKIRQAFPVSPBDWLIAFDAQSBLRV 661

Qy 591 RAGVSSESDVETIE-----MSVADMKE---ASKMILAGVATGLISLFSQ 634

Db 662 LAHTAB---DDNLTIEAFRRGLDHTKTMDIFHVSEEDVTANRROQAKAVNFGIVGIRD 718

Qy 635 KUFLKSSSFORKDMVSMESDVATIGSVRADDESEALPRMDARTAEMIVSKKKQKIKSLAP 694

Db 719 -YGLAQUNLITKREAAFFIERYFASPCVK-----QYMDNIQOE-AKQKCYT 764

Qy 695 GDPVTRIEMLPEDVGRM-LKIWDRAA-ETAQIGLGLVVDYTLLKLSD-VSTVSDGTR- 750

Db 765 TULLRERYLPDTSRNTRVRSFARTAMNTPLQSAAD-IKKAMIDSVRLEBERLQR 823

Qy 751 -----LVENTLESACSLDVLHPE 769

Db 824 LLIQVQHDLILESHPKESIERLCRLV-PE 850

us-09-135-642-32

Sequence 32, Application US/09135642

Patent No. 606483

GENERAL INFORMATION:

APPLICANT: TUDOR, STARLA D.

APPLICANT: RIGGS, MICHAEL G.

TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHERMOPHILUS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Gen-Probe Incorporated

STREET: 9880 Campus Point Drive

CITY: San Diego

STATE: CA

COUNTRY: USA

ZIP: 92121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/135,642

FILING DATE:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/394,232

FILING DATE:

APPLICATION NUMBER: 08/222,612

FILING DATE:

APPLICATION NUMBER: 16-SEP-1994

FILING DATE:

APPLICATION NUMBER: 08/222,612

ATTORNEY/AGENT INFORMATION:

NAME: Fisher, Carlos A.

REGISTRATION NUMBER: 36,510

REFERENCE DOCKET NUMBER: GPB4003.CP2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-533-2807

TELEFAX: 619-546-7929

TELEX:

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 876 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

US-09-135-642-32

Query Match

Best Local Similarity 3.2%; Score 128.5; DB 3; Length 876;

Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

Qy 99 HFLTDGIRRRAFERARVKPPQPGFSPDALISRQLQACETLSNPRREY-NEGLIDDE 157

Db 107 RYEDDITGMARAERB--GFPAVKVSGRDLTQ---LASPQTVETTRKGIDIE 158

Qy 158 EATVITDV-----PMWKVPG-----ALCVIQEGGETEIVURVG 190

Db 159 SYTPETVEKVLTPEQIVDPLKGMDKSDNPNGVPGIGEKTAVKLJKQFGTENVIASI 218

Qy 191 EALKERKUPKSFRQ--DVTLMVALAFLDVSRLA--MALDPDFITGYEPVERALKIQLQE 245

Db 219 DEIKGEKULKENLRQYRDIAL-LSKQLAICRDAPVELTD--DIVKGEBDKVALFQE 275

Qy 246 EGASSLAPPDRQIDE-----TLEBITPRVYLELGLI--PLGDDYAAKRUNG- 290

Db 276 LQPOSFLDRMAYQVOTDEGEEKPLAGMDFAATADSVIDEMLAQDAALVVEVGDNYHAPIVGI 335

QY 291 -----LSGVNLTWSVGGGASALVSGLITREKFMNAPFLRMTAE--- 330  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 336 ALANERGRPFLRPTETALLADPKFLAW-----LGDETKKTMFDSKRAAVALKWGI 385  
 QY 331 -----QVDLFVATPSNIPAES-----FENVTEVALVALQAQPFLGK-----KPHLQ 370  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 386 BLRGVVFDILLAAYLDPQAQDVAVAKMHOYEAVRSDEAVYKGAKRTVDEPTLAE 445  
 QY 371 DADKQFOOLQOAKWMAMEIPAMLYDRTRNWE-IDFGLERGLCILIG----KVDECRM 423  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 446 HILAR----KAAMIWALEPLMDLRNEQDLRTEBQPLAGILANBFTCWKVDTKRL 500  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 424 -WGLD-SEDSOYRNPAVTEVFLENNSRDNDLPGLKLETFLAGVWPRFRDTKOK 481  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 501 BOMGAELTEQLO----AVERRIYELAGOEFNINSP--KQLGT----VLFDKLQLPVLK 549  
 QY 482 PKLGDYDPMVSYL-ERVEVYOGSPLAAAMTMRGAHVKASAMAOLOKVPFSRTD 540  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 550 TKTG-YSTSADVLEKLAHPHETE--HILHYROGKLOSTYB-GLUKVWHPVTGKHTM 605  
 QY 541 RNSAEPKDVQET--VSUDPVGNVGRGEPEGPVIAEVRPSE----NFENDYAI 590  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 606 FNQD----LTQTCRSSLSEPNLONIPRLERKIRKAFAVSPRPDWLJFAADYSQILRV 661  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 591 RACVSESSVDETVE-----MSVADMKE--ASVKAAGVAGLISLSFSQ 634  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 662 LAHIAE--DDNLIEAFRRGLDHTKAMDIFVSEEDVTANMRQAKAVNGIVGJSD 718  
 QY 635 KYFLKSSSFQRDMVSMESDVATICSVRADDSEALPRMDARTAENIVSKWQIKSLAP 694  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 719 -YGLAQANINTRKAEASPIERYVAPFCYK-----QYMDNTVQE-AKQGKVY 764  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 695 GDPHRIMPLEPDIGRM-LKIWIDRAA-ETAOGLVYDYLKLSDVTSADGTRA- 750  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 765 TLMHRYVLPDTTSRNFRSPTAANTPTQGSAAD-LIKAMIDSVRABERIQLR 823  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 751 -----LVENTLEESACLSLQHPE 769  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 824 LLLQVHDLILEKPKEEIERLCRLV-PB 850  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 RESULT 11  
 US-09-135-642-34  
 Sequence 34 Application US/09135642  
 Patent No. 606483  
 GENERAL INFORMATION:  
 APPLICANT: RIGGS, MICHAEL G.  
 APPLICANT: SIVARAM, NATHOR  
 APPLICANT: TUDOR, STARLA D.  
 TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS  
 TITLE OF INVENTION: STEAROTHERMOPHILUS  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Gen-Probe Incorporated  
 STREET: 9880 Campus Point Drive  
 CITY: San Diego  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92121  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/135, 642  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/394, 232  
 APPLICATION NUMBER: 08/307, 410

FILING DATE: 16-SEP-1994  
 APPLICATION NUMBER: 08/222, 612  
 FILING DATE: 16-SEP-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fisher, Carlos A  
 REGISTRATION NUMBER: 36-GP54003.CP2  
 REFERENCE/DOCKET NUMBER: 36-GP54003.CP2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619-533-2807  
 TELEX: 619-546-7929  
 INFORMATION FOR SEQ ID NO: 34:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 876 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULAR TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: internal  
 ORIGINAL SOURCE:  
 US-09-135-642-34

Query Match 3.2% Score 128.5; DB 3; Length 876;  
 Best Local Similarity 21.0%; Pred. No. 0.0053; Gaps 43;  
 Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

Db 99 HFLLDGIRRAREVSKPPQFGSDDAISRROILOQACTUSNPRSRREY-NEGLLDE 157  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 107 HYEADDIGTMARAERB--GFAVKUTSQRDLTQ----LASPQTYEETIKGIDIE 158  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 158 EATWITDV-----PWDKVFG-----ALCVUQBGETEIVARVG 190  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 159 SYTPETVVKYGLTPBQIVDVKLGMGDKSNDIPGVPGIGEKTAVKLKKQFGTVENLASI 218  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 191 EALKERLKPSKQ--DYLWVMAFLDSRVA--MALDPDFITCEYEFVAKLQE 245  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 219 DEIKGEKUKENLKRQYRDLAL-LSKQLAACIDRAPVELT--DIVYKGBDRBKVALFOB 275  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 246 EGASSLAPDLRQIDE-----TLEETIPRYVLLGLI--PLGDDYAKRKG- 290  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 276 LGQSFLDPMAYQTDGSKPLAGMDTADSYTDEMADKAALWVETVGDNTTHAPTLG 335  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 291 -----LSGVNLTWSVGGGASALVSGLITREKFMNAPFLRMTAE--- 330  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 336 ALANERGRPFLRPTETALLADPKFLAW-----LGDETKKTMFD SKRAAVALKWGI 385  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 386 BLRGVVFDILLAAYLDPQAQDVAVAKMHOYEAVRSDEAVYKGAKRTVDEPTLAE 445  
 QY 371 DADKQFOOLQOAKWMAMEIPAMLYDRTRNWE-IDFGLERGLCILIG----KVDECRM 423  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 446 HILAR----KAAMIWALEPLMDLRNEQDLRTEBQPLAGILANBFTCWKVDTKRL 500  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 424 -WGLD-SEDSOYRNPAVTEVFLENNSRDNDLPGLKLETFLAGVWPRFRDTKOK 481  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 501 BOMGAELTEQLO----AVERRIYELAGOEFNINSP--KQLGT----VLFDKLQLPVLK 549  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 606 FNQD----LTQTCRSSLSEPNLONIPRLERKIRKAFAVSPRPDWLJFAADYSQILRV 661  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 591 RACVSESSVDETVE-----MSVADMKE--ASVKAAGVAGLISLSFSQ 634  
 QY ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ; : ;  
 Db 662 LAHIAE--DDNLIEAFRRGLDHTKAMDIFVSEEDVTANMRQAKAVNGIVGJSD 718  
 QY 635 KYFLKSSSFQRDMVSMESDVATICSVRADDSEALPRMDARTAENIVSKWQIKSLAP 694

Db 719 -YGLAQNLTIRKKEAIFIERYFASPGVK-----QYMDNIVQE-AKQGYT 764  
 Db 695 GPDHRTMLPEVLDGRM-LKIWIDRRA-EIAQOLGLVYDYLKLSDV-SVTVSADGTRA- 750  
 Db 765 TLRHRRYLPDTTSRNVRSAERTAANTPTQGSAA-D-IKKAMIDSVRRLERLQAR 823  
 Qy 751 -----LVEATLESACISDLVHPE 769  
 Db 824 LLIQVHDBLILEAPKERIERLCRLV-PE 850  
 RESULT 12  
 US-08-394-232A-20  
 ; Sequence 20, Application US/08394232A  
 ; Patent No. 6100078  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RIGGS, MICHAEL G.  
 ; APPLICANT: SIVARAM, MATHOOR  
 ; APPLICANT: TUDOR, STARLA D.  
 TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHERMOPHILU  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Gen-Probe Incorporated  
 STREET: 9880 Campus Point Drive  
 STATE: CA  
 COUNTRY: USA  
 IP: 92121  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/3941,232A  
 FILING DATE:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/307,410  
 FILING DATE: 15-SEP-1994  
 APPLICATION NUMBER: 08/222,612  
 FILING DATE: 16-SEP-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fiber, Carlos A.  
 REGISTRATION NUMBER: 36,510  
 REFERENCE/DOCKET NUMBER: GP94003.CP2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619-531-2807  
 TELEX:  
 TELEFAX: 619-546-7929  
 IN INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 876 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: internal  
 ORIGINAL SOURCE:  
 US-08-394-232A-20  
 RESULT 13  
 US-08-394-232A-32  
 ; Sequence 32, Application US/08394232A  
 ; Patent No. 6100078  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RIGGS, MICHAEL G.  
 ; APPLICANT: TUDOR, STARLA D.  
 ; APPLICANT: SIVARAM, MATHOOR  
 TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHERMOPHILU  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Gen-Probe Incorporated  
 STREET: 9880 Campus Point Drive  
 CITY: San Diego  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92121  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 Db 159 SVTPETVKEGYLTPEQIVDVLKGIMGDKSDDNPVGPGCIGKPAKUKQFGAVENVLASI 218  
 Db 191 EALKERULPKSKPQ---SVWVVALAFPLDVSDRA---MALDPDPDINGEYFEAKLQLQ 245  
 Db 219 DEKGKEKLNKRQYRDAL-LSKQLAACIRCAPEVTLTD-DIVKGBDRKVWLFQE 275  
 Qy 246 EGASSLAPDLRAQIDE-----TEETPRVYLELGLI--PLGDXAAKRLNG- 290  
 Db 276 LGFOSPLDKMAYQTDGEGKPLAGMDFATADSDTDEMADKAALUVVUGDNYHAPIVG 335  
 Db 291 -----LSGVRNLWSVGGGASALVGGLTREKFMNEAFLRMTAE--- 330  
 Db 336 ALANERGRPFURPETALADPKELAW-----LGDETKKTMDSKRAAVLKWKGI 385  
 Qy 331 -----QVDIFVATPSNIPAES-----FEVYEVALLVAQDFIGK-----KPHLLO 370  
 Db 386 ELRGVVVFDLILAYLLDPAQAGDVAAVAKOMQYEAVRSDEAVYGKAKRTVPDPLAE 445  
 Qy 371 DADKQFQOLQQAKUMAMBI PAMYDTRNME-IDFGLERGLICALLIG----KUDCRM 423  
 Db 446 HLAR-----KAATAWLEBPLMDRNEODRLTLEBOPLAGILANMFCTGVVKDPLR 500  
 Db 424 -WGLD-SBDSQVRNPAIVEFLVLENSRDDNDLPLGCKLLETFWLAGVVFPRFDTKDK 481  
 Db 501 EQMGAEITEQLQ-----AVERRYELAQEFNINSP---KQLG-----VLFDKLQPLVTK 549  
 Qy 482 FKLGDDYDPMWISYL-ERVEVQGSPLAAATMARTGAHEVYKASAMQALQKVFPSPRYD 540  
 Db 550 TKTG-YTSADVILEKPLAHPHEVE-HILHYRQLGKLOSTIE-GLIKVHPVTGKHTM 605  
 Qy 541 RNSAEPKDQET--VFSVDPVGNNGRDGECPVIAEVRSE-----NFETNDYA 590  
 Db 606 FNQA---LTGTRGLSSVPEVNQNTPIRLEBGRKIROAFVFPSEPDWLFIAADYSQIBLRV 661  
 Db 591 RAGVSSESDVETE-----MSVADMKE--ASVKAAGVAGLISLFSQ 634  
 Db 662 LAHTAE---DDMLIAFRRGDQDHTKAMDFHVSEBDVTTMRQAKAVNFGIVVGSID 718  
 Db 635 KYFLKSSSSFORKDVMWSMMSDVATGVSADDSEALPRMDARTAENIVSKWQKIKSLF 694  
 Qy 719 -YGLAQNLTIRKKEAIFIERYFASPGVK-----QYMDNIVQE-AKQGYT 764  
 Db 695 GPDHRTMLPEVLDGRM-LKIWIDRRA-EIAQOLGLVYDYLKLSDV-SVTVSADGTRA- 750  
 Db 765 TLRHRRYLPDTTSRNVRSAERTAANTPTQGSAA-D-IKKAMIDSVRRLERLQAR 823  
 Qy 751 -----LVEATLESACISDLVHPE 769  
 Db 824 LLIQVHDBLILEAPKERIERLCRLV-PE 850



Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

APPLICATION NUMBER: PCT/US95/04080  
FILING DATE:

CLASSIFICATION: INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS: LENGTH: 876 amino acids

TYPE: amino acid  
TOPOLGY: linear

MOLECULE TYPE: protein

PCT-US95-04080-20

Query Match 3.2%; Score 128.5; DB 5; Length 876;  
Best Local Similarity 21.0%; Pred. No. 0.0053; Mismatches 302; Indels 201; Gaps 43;  
Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

QY 99 HFLTDGIRRRAFRARVKSPQFGRFSDDALISRSQILQACETLNSPRSRREY-NEGILIDDE 157

Db 107 HYEADDIGTMARAAERB--GFAVKVIGDRDITO---LASPOVTVEITKKGTIDIE 158

QY 158 EATVITYV------PMDKPG-----ALCVIQLQEGGETEVIRVG 190

Db 159 SYTPETVVEKYGLTPEQIVDLCGKLMGDKSDNITPGVPGIGEKTAVKLIKQFGTVENVLASI 218

QY 191 EALKERLIPKSQK-DVVLVMAFLDVSDRA--MALDPDFITGYEFVBALKIQE 245

Db 219 DEIKGEKLUKENRQYRDLAL-LSKQLAACRPAVELTD--DIVVKGEDRKVVALFQE 275

QY 246 EGSSLAPLDLRAQIDE-----TLEETTPRYVTELLGL--PLGDDYAAKRLNG- 290

Db 276 LGFQSFLLDKMAYQVTDGEGKPLAGMDFAIDSVDIEMLAADKAALVVEVGDNTHAPIVGI 335

QY 291 -----LSGVNLWLSVGGGASALVGLTRKFMNEAFLMTAAB--- 330

Db 336 ALANERGRFFLRPETALADPKFLAW-----LGDETKKTMFDSKRAAVALKWKGI 385

QY 331 -----QVDLFVATPSNIAES-----FEVTEVALVALVAQAPIK-----KPHLQ 370

Db 386 ERLGVVFULLAYLDPQAQDVAVAKHOMYEAVRSDEAVYGKAKRTVDEPTLA 445

QY 371 DADKQFOOLQQAKUMAMBIPIAMYLYTRNNWE-IDFGILERGLCALLIG----KVDCRM 423

Db 446 HLAR-----KAAAIWALEBPLMDELRRNEQDRLTLEQPLAGILAMNEFTGVKVTKL 500

QY 424 -WLGLD-SEDSQRNPNAIVEFVLENSNRDDIDLPGLKLTETWLAGVVFPRDTKDK 481

Db 501 EONGAELTEQLO---AVERRYYELAGQEFNINSP---KQLG---VLFDKLQLPVLK 549

QY 482 FKLDGYDDPMVLSYL-ERVENQGSPLAAATMARIAGAHTVKASAMQALQKVPFSYTD 540

Db 550 TKIG-YSTSADYLEKLAAPHETE--HILHYRGQLKQSTYB-GLIKVWHPVTGKHTM 605

QY 541 RNSAEPKVQET--VFSVDPVGNVGRDGEPGVFIAEVRPSE-----NFETNDYAI 590

Db 606 FNGA---LTQTCRLLSVEPNLQONIPRLEGKIROAFVPSBDMLIFADYSQIELRV 661

QY 591 RAGVSESSVDETVE-----MSVADMLES---ASVKILAGVAGLISLFSQ 634

Db 662 LAHIAE---DDNLLIEAFRRGDLIHTKTAMDIFHVSEEDVTAMRQAKAVNGIVYGISD 718

QY 635 KTFLKSSSSPKDMVSMESVATIGSVRADSEALPRMDARTAENIVSKWQKIKSLAP 694

Db 719 -YGLAQONIINTKREAAFPYFASFPGVK-----QYMDNTIVQS-AKOKGVT 764

QY 695 GPDHRRIEMLPEVLDGRM-LKIWTDRAA-ETAQGLVVDYTLIKLUSD-SVTVSADGTRA- 750

Db 765 TLIHRRVLPDITSRNFRVSPERTAMNTPIQGSAAD-IKKAMIDLSVRLERLQAR 823

QY 751 -----LVEATLBBEACSLDLVHPE 769

Db 824 LLLQVHDBLILEAPKEEIERLCLRV-PE 850

RESULT 15  
PCT-US95-04080-20

; Sequence 20, Application PC/TUS9504080  
; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS

; TITLE OF INVENTION: STEAROTHERMOPHILUS

; NUMBER OF SEQUENCES: 34

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0., Version #1.30 (EPO)

; CURRENT APPLICATION DATA:  
; Search completed: June 10, 2005, 01:42:43

, Fri Jun 10 09:57:54 2005

us-10-600-070-2.rai

Page 13

Job time : 33 secs

THIS PAGE BLANK (USPTO)



123 PNPPPGGLPQDYLISVILAHWELSRERWQQQ-----YFPAATSKALARLQQDNDFP 176  
 Db  
 Qy  
 254 DLRAQIDETLEETTPRYVRLGLGPP-LGDDYAAKKLNGISGVNLWLSVGG-GGASALVG 311  
 Db  
 Qy  
 177 ALEARIRQELYRLRPRYRLLAKEBGGGER--ORQOGLALLOQANVODRGGIEGKQBDVS 233  
 Db  
 Qy  
 312 GLTRE--KFMEARFLRMTAEGQDLFVATPSNPAESFSEVTEVALVALVAQAPIKKPHL 368  
 C:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
 C:Accession: AB2494  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Yano, A.; Iriuchi, Nakazaki, N.; Shimojo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AB2494  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2297 <KUR>  
 A:Cross-references: UNIPROT:QBYL08; GB:BA000020; PIDN:BAB78214.1; PIB:917135668; GSPDB:GCA  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: al17130  
 A:Genome: plasmid

QY	149 YME-----GILDDREATVITDVPDKVKPVALCVCVQRGGEETIVARVEALKER 197
Db	312 FISVENITSLINLALLDKDROBERTFWFQQALBALPEKTI---ENGVN--IFSIAPQLTER 365
Qy	198 LPKSFQKVYVLMALAFLLVSDRMDP----DFITGYFVEALKLQER---G 247
Db	426 LAIAIPHJKETQCNVLYVILVERENIWPFI-----SGVSETLV 465
Qy	366 BPLLEKTEBLVNNLERGYOREANAVATAPYLKSEPILOQALRATIAYKIANEDRIKA 425
Qy	300 SYVGGGAGAALVGGTNTREKINNEATIRMTAEQVDIFVATPSNIPRE-SFVVTEVALVA 358
Db	466 FI---ASYL--SKSOSKLUQKAR-----NIVONLFRGFSYDQIALVALA 502
Qy	359 QA-FIGK-KPHLQDADKOFQOLQO--AKUMAM-----EIPAMYDTRNWEI 402
Db	503 LAPHLKSLPQLQALQKAINTEGASKVPAVIFPHLPOSQSOLLEKAFLIQTIE 562
Qy	440 VERVLE---NSNRDDNDLPGIC-----KLIETMLAGVVFPRFRDKFKEKLG 486
Db	622 LQKALBIAKKINSDEKOBALAAIAYOLSESEPBILLEQVLTQKTAQTYHSHHKRITAIIP 681
Qy	487 YVDDPMWLSYLERVEVVOGSPLAAMATMARIAGAEMVKASAMQ-AUQKVPSRYDORNSE 545
Db	682 VLRSPQQTPELQWVFSATASSLAAAPCL--PEPKRSVLOQALKMIDANYSULRAN 738
Qy	546 PKDVQTVTFSVDPVGNNYGRDGERGVFIAEVRSNFENDYATRAGVSESSVETV- 604
Db	739 ELKIVIPLSKT-----EQEKVIEIAKTK-NDFKGAELLAFLVATHLSEFFQVKIP 790
Qy	605 ENSVADMK--EASKVLAAGVAILGLISFSQXFLKSSSSFKRDMSMESVATIC 661
Db	791 PBLIERIKAIENDUSKAKALUVVYPLSKSSPESLDKAFEAENLQYQSCQFDPLVTLA 850
Qy	662 S-VRADDSBALPRMDARTMENTVSKWOKTSLA 693
Db	851 THLKRECETKLLQALEKAKDIDSEHQADDA 883
RESULT 4	
Qy	A48757 acetyl-CoA carboxylase (EC 6.4.1.2) - Cyclotella cryptica
C;Species: Cyclotella cryptica	
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004	
C;Accession: A48757	
R;Rossiter, P.G.; Ohlrogge, J.B.	
J; Biol. Chem. 268, 19254-19259, 1993	
A;Title: Cloning and characterization of the gene that encodes acetyl-coenzyme A carboxylase. Accession number: A48757; MUID:93374903; PMID:8103514	
A;Status: Preliminary; nucleic acid sequence not shown	
A;Residue: 1-ROB>	
A;Cross-references: UNIPROT:Q39478; GB:L20784; NID:91065903; PID:AAA81471.1; PID:940945	
A;Note: authors translated the codon GCC for residue 1834 as Ala	
C;Genetic:	
A;Introns: 25/1; 729/1	
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-C;Keywords: biotin binding; ligase	
P;99-603/Domain: biotin carboxylase homology <BCH>	
P;731-803/Domain: lipoyl/biotin-binding homology <LPB>	
P;770/Binding site: biotin (Lys) (covalent) #status predicted	
Query Match 3.5%; Score 143; DB 1; Length 2089; Best Local Similarity 20.5%; Pred. No. 1-8; 289; Indels 272; Gaps 44; Matches 169; Conservative 113; Mismatches 907	
RESULT 5	
Qy	JC4286 DNA-directed DNA polymerase (EC 2.7.7.7) - Bacillus stearothermophilus
N;Alternative names: DNA polymerase I	
C;Species: Bacillus stearothermophilus	
C;Date: 14-Nov-1995 #Sequence_revision 08-Feb-1996 #text_change 17-Mar-1999	
C;Accession: JC4286	
R;Phang, S.M.; Teo, C.Y.; Lo, E.; Wong, V.W.	
Gene 163, 65-68, 1995	
A;Title: Cloning and complete sequence of the DNA polymerase-encoding gene (Btpoli) and Reference number: JC4286; MUID:9601245; PMID:755480	

A;Molecule type: DNA  
A;Residues: 1-879 <PDB>  
A;Cross-references: GB:U23149  
A;Note: The authors translated the codon GCG for residue 15 as Arg  
C;Comment: It has three enzymatic functions: 5'-3' exonuclease activity, DNA synthesis  
C;Genetics:  
A;Gene: BSTRPOLI  
C;Superfamily: DNA-directed DNA polymerase I  
C;Keywords: nucleotidyltransferase

similar to ch-TOG protein from *Homo sapiens* [imported] - *Arabidopsis thaliana* [Species: *Arabidopsis thaliana* (mouse-ear cress)]  
C-Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C-Accesion: AB4771  
Rlin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M., Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
eub, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.  
*Nature* 402, 761-768, 1999  
A1: Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A2: Reference number: AB4420; MUID:2008487; PMID:10617197

	Best Match	Local Similarity	Conservative	Pred.	No. of Mismatches	Indels	Gaps
Oy	99 HELTDGIRRAFEARSKPQFGSDDALISRQLQACETLSNPRSRREY-NEGLDDE	22.0%	22.0%	22.0%	175	132	44
Db	108 HYEADDIIGTMMARAEI--GFAVKVSGDRDLTQ----LASPVQTVETIKRGIDIE	157	157	157	157	157	159

A>Status: preliminary  
A:Molecule type: DNA  
A:Ridues: 1-2021 <S>  
A:Cross-references: UNIPROT:Q9ZQN6; GB:AE002093; NID:94263790; PIDN:AD15450.1; GSPDB:GN  
A:Genes: At2g35630  
A:Map position: 2

Db	SYTPETVVEKYGLTPEQTVDLKGLMGDSDNIPGPGIGKTKAVKLKQFGTVENVIASI	219
Qy	191 EALKKERUPSKPSHQ--DVTLLWALAFLLDSRDA--MALDPDPPTIGYEFVEALKLQE	245
Db	220 DEKGKEKUNERQYRDIAL-LSKQLAICRDPVLTID-DIVVKGEDRKVVALFQE	276
Qy	246 EGASSLAIDLRAQIDE-----TLEETTPRYVLELGI--PLGDDYAAKRLNG-	290
Db	277 LAGPSFLDKMAYQTDLEGKPLAGMDPATAFADSVDYDMLMADKAALVUEVGDNTHAPIVGI	336
Qy	291 -LSQVRNTIW-----SVGGGASALVGSLITREKFM-----NEAFLRMTAAECQY-D	333
Db	337 ALLANERGRFRFLRPETAVADPKELAWLGDETKKKTMFDSSKRAAVALNSKGIELAGGVVF	396
Qy	334 LEVATPSNIPAES-----FEYEVALALVMAQFIGK-KPHLQDADKQFQQL--QQAK	383
Db	397 LLLAAVLLDPAQAGDVAVAKNHOYEAVRSDEAVYVGAKRRTVDPDETLABOLVNRQAA	456
Qy	384 VMMAMEIIPAMLYDTRNNW-IDFGLERGLCALLIG----KUDECRM-WLGLD-SEDSQY	434
Db	457 IWALEPLMDLRRNEDQRDTTEHALAGILANMEFTGVKUDTKRUEOMGAELTSOLQ-	515
Qy	435 RNAPIAVEVLEMSNRDDNDLDPGLCKKILETWTLAGVVPRFRDTCDDKPKFLGJYDDDMNL	494
Qy	516 --AVERRIYELAGOEFINSP--KQLTG--VLFKDQLQPVLKKTGT-YSTBADVL	564
Db	495 SYL-ERYEVWVOOSPLAAMATMARIGAEIWKASAMOAJQKVFSRYSRTNSAEPKDQBT-	552
Db	565 EKLAPHRIVE--HILHYRQLGKLUQSTY-GLUKVVHPVTKVHMFNOA---IQTG	617
Qy	553 -VFSVPDVGNVNGORDGPGCVFTAEEWRPSE-----NPETNDYAIRAGVSBSSEBDTT	603
Db	618 RLSSVEENTLQNIPITRLEBGRKIRQAFVPSERPDWLIAFDYQSIELRVLAHTAE--DDNL	674
Qy	604 VE-----MSVADMILKE--ASKVILAGVAIGLISLFFSOKYFLKSSSFORK	647
Db	675 IEAFRMRWLDIHTKTANDIFHVSEBDVTAMRROAKAVNFGIVYGISD-YGLAQNLWTRK	733
Qy	648 DMVSSMMSDVATIGSVRADDSBALPDRDARTAENIVSKWQKTKSLARGPDRHITEMPEVL	707
Db	734 EAFAEFTIERYFASPPGVK-----QYMDNIVOB-AKOKGYVTLHRRYLPDIT	780
Qy	708 DGRM-LKLTWIDRA-ETRQLGLVYDYLKLSVD-STVSVADGTRA-----LVEA	754
Db	781 SENFNFTVTAERTAMNTPIQGSAD-IKKAMIDSVSREERLQARLLOQHDELLEA	839
Qy	755 TLEESACSLDLYHE	769
Db	840 PKEEIGRLCRLV-PE	853

Query	Match	Score	DB	Length	2021;
QY	39 ASKWARDLISDFNFTSDSSSFATTTATLVSLLPSIDRPERAVIPIDFYQVGAQT	3.4%	DB 2;	2021;	Best Local Similarity 20.2%; Pred. No. 3.6; Mismatches 179; Conservative 222; MisMatches 315; Indels 272; Gaps 43;
Db	297 ATKMSERKEVAEUDTKLAASTKKLAEGDFSEBICTRUKKLID-TDVNLAVAVEAIQAGIN- -	98			352
QY	99 HFLLTDGIRRAFEARVSKPQPGFESDDALISRQIQIAACETLSNPRSRRENEYGLJDEE	158			
Db	353 -LACGLKTHFSA-----SSRFMPLVLEKU-----KSKKQSTDPLT	388			
QY	159 ATVITDVPWDKVFGALCVLQEGERETIVRVGEALLKEULPKSFQDVWVMA-----LA	213			
Db	389 QTLQT-----MYKAG-----CLNLVDVTEGKNTAVINKVPLVRSSLTWL	431			
QY	214 P-LDVSRDAMALDPPDFITGYEFEEBAULKQLEGGASSLAPDLR-----AQID-----	260			
Db	432 FCLETSNKALILK----AHKEYVFLCMCBLNDG-----TPDVDRBAAFSALAIAKSVMG	481			
QY	261 -----ETLEITTPRYVIELLGLFLGDDYAAKRNLNGLSGVNNTIWSVGGGASALVGL	313			
Db	482 RPLERSLEKDDYRKKKLSEMIAQGGDQA-----GSSV-TVQSSVGSTATGCLHNRY	535			
QY	314 TREKFMM--NEATLRTMMAE-----QVNLIVFAT-----PSNI	342			
Db	536 TISKPMQNSDASKVFKSASMLSQRPAQASKVGKPGGGKKDGSVRNEGSKSVE	595			
QY	343 PAESFEEVVALVALVAQAFIGK-KPHLQDADQP-----GOLQQAXUMAMELPAMLYDTRN	398			
Db	596 PPDVEPAREMGLEBENR-LGSLVTKPETYVSOLKSSVWKERLEATLALKIEGL-----648				
QY	399 NWELDFGLE--RGICALLIGKDVCECRMWNLGDSQRNPAIVEPVLEN---SNRDD-	451			
Db	649 -QLDKSVELLVLLCAV-----PGWNKVNQL-----FSEENYFFPSNLPDF	689			
QY	452 -----NDLPLGCKLKLTWTLAGVVFPPFRDTKDKKKFKLQGDYDYPDMVL	494			
Db	690 SINSLGTTSERVADIKTRASAMKCUTAFCEAVGPGFVERL-----FKIMKEHKNPKVL	742			
QY	495 S-----YLERVEVVOGSPLAAATMARIAGEH-----521				
Db	743 SEGILWWMTSAVDDPGFVSLJKLKLIDDFCKDVGJQSSTAATRNATIKLGLGALKHFVGPDIK	802			
QY	522 -----VKASAMQALQKVFPDSRYTDRNSAEPKDVOBTVFSVPVGNNVRGRDGBGPWFIAEA	576			
Db	803 GFINDVKPALKSALDTEY-EKNPFFGTAAPKRVWT-----SVSTSISGGDSLPRDISTK	859			
QY	577 VRRS--EUFETNDYAIRAGVSESSVDETTVEMSVADMILKEASVKKLAAGSYAIGLISLPSQ	634			
Db	860 ITPNLKGPFESPDMKRL---ESIE-----ANVKILBANKRJQPTGIG-----ELFGG	905			
QY	635 KYTIKSSSFORKDMVMSMESDVATIGSVRADSEALPMDARTENIV-----SKWQK	688			

**RESULT 7**

DJBE28  
DNA-directed DNA polymerase (EC 2.7.7.7) - human herpesvirus 3  
C;Species: human herpesvirus 3, varicella-zoster virus  
C;Date: 30-Sep-1988 #sequence\_revision 30-Sep-1988 #text\_change 09-Jul-2004  
C;Accession: B27214  
R;Gen. Virol. 67, 1759-1816, 1986  
A;Title: The complete DNA sequence of varicella-zoster virus.  
A;Reference number: A27345; MUID:86306657; PMID:3018124  
A;Accession: B27214  
A;Molecule type: DNA  
A;Cross-references: UNIPROT:P09252; ENBL:X04370; NID:95989; PIDN:CNA27911.1; PID:960017  
C;Genetics:  
A;Gene: 28  
C;Superfamily: DNA polymerase  
Query Match 3.2\*; Score 131.5; DB 1; Length 1194;  
Best Local Similarity 19.0%; Pred. No. 41; Mismatches 168; Conservative 122; Mismatches 297; Gaps 44;  
Db 48 SDDNFTSDSSS--SSFATATTATLVSLLPPSTDPERHV--PIPIDTYQVLRQ---TH 99  
205 NDATLNGDKNAFFGTSFSKSASPPSPRVEV--IERTDVTYDTPQPCATYRVSPSSKETN 261  
100 PLTGDGIR--RAREARVKKPQPGPSDALIS-----RQI---- 132  
Db 262 YLCDDNFHPELKYYKEGRVDAATRFLMDNPQFGVSVFQWYQLKPGVQGVGERVRVRPASRQLTSD 321  
Qy 133 LQMAC---ETISNPRSRRTVEGLLDE-----EATVITDVPMKVKPGALC 175  
Db 322 VEDCMDSNLQAFPLNDSPWDVYKLCFCDFECKSGGSNSLAPPTHDLD---VIQJSC 377  
Qy 176 VI----QEGGETTIVLVRGEBAHLKERLPRSKFKO-----DVAUVMALFLVDNSRD 220  
Db 378 LYSIPROSLHEHILPSLGCDIPQRYQEMKAGLPPTVLFDPSEFELLIAFMILVKQ 437  
Qy 221 AMALDPDPFPTGSEPERVERAKLKLQBEGLASSLADPLRAGIDETEKEITERYVMLGLPLG 280  
Db 438 Y---APEFATGYNIVNFDWAFT-----MEKLNSIYSLKL----- 468  
Qy 281 DDYAAKRKLNGLSVNRNLMWSGGCGASALVGIGTREKPKNAFLRMTAQDVLVATPS 340  
Db 469 DGYGSTINGGLFKI---WDVKGSG-----FQRSSKVKINGLISLNUVATE 512  
Qy 341 NIDAESFENVEA-----LAIVAQATG 363  
Db 513 KULSSSYKLDVSRAREALMWSKRLPKYKIPGVIASGPTRGTCYECYQDPSALVGKLFK 572  
Qy 364 KKHLLOPDADKQFOOLQOKKVMAMEIPAMLYDPTTRNNWBDGJLFRGLCALLKQVDCRM 423  
Db 573 YLPHLELSA-----VARLARITLTKAID-----GOVRYTCLGLASS--- 612  
Qy 424 WLGDSLSEPSQYRIPAIVER-----VLENNSRDDNDLPLCILKULETMAGV--FPR 473  
Db 613 -RCPFLPQGY-PATPEVKDVTIDPGVDEEEEDBESVSPCTGSSGRNGVYKGARVED- 668  
Qy 474 FRTDKDKUPKLGQYDDEPNLMSLIERV-EVVOQSPPLAAAATMARGAERHKASAMQLQ 532  
Db 669 -PDT-----GTVIDPVVWLDRASLYPSIIQAHNL--CFTTITLNFTVK-----R 710

**RESULT 8**

A70668  
mycocerosate synthase (EC 2.3.1.111) - Mycobacterium tuberculosis  
C;Species: Mycobacterium tuberculosis  
C;Accession: A70668  
R;Role: S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feil, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: A70668  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Cross-references: UNIPROT:P96291; GB:Z83858; GB:AU123456; NID:93261675; PIDN:CAB06108.1  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: mas  
C;Superfamily: mycocerosic acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I homolog; acyl homology; [acyl-carrier-protein] S-malonyltransferase homology  
C;Keywords: acyl transferase; carrier protein; coenzyme A; phosphotransferase  
F;17-426;Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OS>  
F;534-815;Domain: [acyl-carrier-protein] S-malonyltransferase homology <ANT>  
F;144-1733;Domain: long-chain alcohol dehydrogenase homology <LDH>  
F;1765-1945;Domain: short-chain alcohol dehydrogenase homology <SDH>  
F;2016-2096;Domain: acyl carrier protein homology <ACP1>  
Query Match 3.2\*; Score 131.5; DB 2; Length 2111;  
Best Local Similarity 19.9%; Pred. No. 10; Mismatches 151; Conservative 82; Mismatches 237; Indels 289; Gaps 33;  
Matches 151; Conservative 82; Mismatches 237; Indels 289; Gaps 33;  
Db 151 EQLIDDEPRTAVTIDVP--WDK-----VPGALCVLQEGCETETIVLVRGEALLKRL 198  
Db 29 ESLURGDD--LVIEBIPPDWDADDYDPEPGVPGR-SVSRWGG----- 68  
Qy 199 PKCFKQDYLWMAFLDVSRSRDAWALDPPFI--TGYEVFUEALKLQIQQEGGSSLAPDL 255  
Db 69 --FLDDYAGFDFAEFGFISREATESDPOQRLLETSWEAISHA----- 109  
Qy 256 RAQIDETLEBRIPRIVLGLPLGDDYAKRKLNGLSCWRNLWMSVGGCGASALVGITR 315  
Db 110 -----GLPDSL-----AGSSTAVFTGITH 129  
Qy 316 EKFPMNEAFRMTAEEQDVLFLVATPSNT-----PAESFVV----- 350  
Db 130 EDYL--VLTTFAGGLSFYVTLGLNNNSVASGRIAHITGLHGPAFTIDACSSGLMAVHL 186  
Qy 351 -----EVALLALVQAFIGKKRPHL-----QDADKQFQQLQQA 383

Db	187	ACSL1HOCRADALAGGCAVLEPHASVYASACQMSSTGRCHSPFADAGEFVRSECGM	246	Qy	341	NPIAESFEVTEVALAYQAFTGKPHLQDANKOFOOLQOKA KUMA-----MEPAM	340
Qy	384	VMAIMEIIPAMLYDPTTRNNWBIDFGLERGLCALLIGKVUDECRMWGLDSPSQYRNPAVEFV	443	Db	581	KUNVSELSNAEKOKLGLESVLUKGVEAQDVLSPAQOOLMQHLYKIMABOTKTICKNDI	640
Db	247	VILKRLPDALDRGNNR---IFAVRTGATNQDGRTET---LTMPSEDQV---AVVRAA	295	Qy	393	LUTVNNUWE-----IDFGLERGLALLIGKVUDECRMWGLDSPSQYRNPAI	439
Qy	444	LENSNRDDNDLFLCKLETWLAGWPRERETDKKFKELGIYDDMVLSLREVY	503	Db	641	LDPBSWELKTNIQITSNWLDPTAEGEILDAITNTAGSLEAQD---KAII	696
Qy	296	LAAG-----VOPETVGVV-----EAHGTTGTPGDPTEYRSILRV-YG	332	Qy	440	VERTLEN--SNDDNDLPLGCKULETMLAGTVPFRDTKDKKKFKGDYDDPMVSYL	497
Db	504	QSSPLAAATMARTIGAEHKASAMQALQKV-----PSRYTDNSAEPKDVQVQTVFS	555	Db	697	IKCGETIATHSDTSLSLPNKALIMASAEGKIA-----ESQTNLPD--RELMTKGLV	746
Qy	333	AGTPCALCSAKSMGHSRASACTVGLKALISRHGVPPHLHNRUDELSV-----	386	Qy	498	ERIEVVOQSPLAAATMARIAGHEHKASAMQALQKV-----FPRSYDNR	542
Db	556	VDPVGNNVGRDGPGRGEFAEAVERPSENRETNDY-A-IRAGVSESSVDETTVENSVADM	614	Db	747	DJYEGKGKGPETKAVSGIDNSINIDESEKALKKADAASBAALDRDTONTLEFGKQN	606
Qy	387	-----ETGLFVPOAVTPWN-----GNDHTPKRVAVSSFGMSGTNHALVEAPAE	432	Qy	543	SAPRK-----DVQETVFSVPVGNNVGRDGPGRGEFVIAE-----AVRP	579
Db	615	MARTAENIVSKWOK-----IKSLARGPDER-----IEMLPVFLGRML	712	Db	807	IERKHPADIVKAREINAVNPVIELEKSEPVSAERI+VOETSSLINNISKAVEK	866
Qy	433	ASAPESSPGDAEVG-----PRFLMSTS-----SDAL-R	461	Qy	580	SEFETNDYAIAGVSSVDTTIVENSVADMILKEASVYLKLAGVAGLISHFSQKTFK	639
Db	674	MARTAENIVSKWOK-----IKSLARGPDER-----IEMLPVFLGRML	712	Db	867	VNNF-----RAMLSPGNLKTYLEK-----KESAKKVE-----LVKAFGTKESTE	908
Qy	462	QTARQLATWVEHQDCVAMSILAYTLAGRAGRHPRTAVANLPVELVEGLRVAQDAL	521	Qy	640	SSSFORDKMS----SNESDATIGSVRADD--SELPROMDARTAE-NIVSKWOKS	692
Db	713	-----KWW-TDRAEATPQLG-----WVYTLKLISVSVSVTSDGITALV	752	Db	909	ERQSPFIKNLIDKTLUSKEVRQDILQLOEQQRKSEAIENPSVKTEDVRYVSGSKKU	968
Qy	522	YDRAVGHGDRGPVWVFSGSOVNMGSTOLLASEPVTAAKL--EPVIAESGSVT	578	Qy	693	A 693	969
Db	753	EA-TLEEACSLVLPHEPATATVTTYTRYFWKSG	790	Db	969	S 969	969
Db	579	EATAQQRTGIDKQVPAVQVQVALATMEOQTYGVRCG	617				
<b>RESULT 9</b>							
C;Species:	Rickettsia conorii	(strain Malish 7)					
C;Accession:	30-Sep-2001	#sequence_revision 30-Sep-2001					
R;Ogata, H.; Audi, S.; Renesto-Audifren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Roc	Science	293, 2093-2098, 2001	#: Evolution in Rickettsia conorii and Rickettsia prowazekii.				
A;Title:	Mechanisms of						
A;Reference:	number: A97700;	MUID:21442074; PMID:11557893					
A;Accession:	C97783						
A;Status:	preliminary						
A;Molecule type:	DNA						
A;Residues:	1-1056 <KOR>						
A;Cross-references:	GB:AE006914; PIDN:AA03205.1; PID:g15619756; GSPDB:GN00173						
C;Genetics:							
C;Gene:	scac4						
<b>Query Match</b>							
Best Local Similarity	3.24;	Score 130.5;	DB 2;	Length 1026;			
Matches	124;	Conservative 109;	Mismatches 277;	Indels 151;	Gaps 25;		
Qy	115	KPPQFGFSDDALISRROTLOQACETLSNPRSRREYNGEQLDDEBEATVITDVFMDKV	170	Db	171	PGALCVQIPEGGETEVILVRGEALLKRLPKSFQDVULMA-LAFLDVSDAMALDPPD	228
Db	378	EPQYKQOQVPPITPTQPLQ--ETSQMPOS-OQVNPNLNL-TATAUSGSMDQIAYVN	432	Qy	433	AGIKAIDSNKQDIDLIKERATAATLNNE-----KSDIAEKOANTIAENTVNKNLKD	487
Qy	229	FITGYEVFEALKLQLQERGASSIAPDLRAQI-----DETLEETPRTYDEL--LGLPLG	280	Db	488	KTGIVNAVLETTKNDQWPNEKSMLETAIVTIVNSNLERPKQQMLEKAVDUGSLIK	547
Db	488	KTGIVNAVLETTKNDQWPNEKSMLETAIVTIVNSNLERPKQQMLEKAVDUGSLIK	547	Qy	281	DYDAKKGKINGLSGVRLTILWSVGGGASALVGGLTREKENNEAPLMTPAEAOUDLFVATPS	340
Qy	548	DD--ASRAAALDQIKDV-----KISNLSPEDKMLAVGD	580	Db	580	QDQDTELEITPRYVTELLGIGPLGDDYAAKELINGLSGVRLTILWSVGGGASALVGGL	315
<b>RESULT 10</b>							
B44110	mycocerosate synthase (EC 2.3.1.111) - Mycobacterium bovis						
C;Species:	Mycobacterium bovis						
C;Date:	05-Dec-1998	#sequence_revision 05-Dec-1998	#text_change 09-Jul-2004				
C;Accession:	B44110						
R;Matnur, M.; Kolattukudy, P.E.	J. Biol. Chem.	267, 19388-19395,	1992				
A;Title:	Molecular cloning and sequencing of the gene for mycocerosic acid synthase, a n						
-Guerin.							
A;Reference:	number: A44110;	MUID:92406887; PMID:1527058					
A;Accession:	B44110						
A;Molecule type:	DNA; protein						
A;Residues:	1-2110 <PAT>						
A;Cross-references:	UNIPROT:Q02251; GB:M95808; NID:9149977; PIDN:AAA25369.1; PID:g149980						
A;Note:	Sequence extracted from NCBI Backbone (NCBI:N13897, NCBI:P113899).						
C;Superfamily:	mycocerosic acid synthase, 3 oxoacyl-[acyl-carrier-protein] synthase I homology						
C;Homology:	3-malonyltransferase homology						
C;Keywords:	acyl-transferase; carrier protein; coenzyme A; phosphotransferase						
r;27-426/Domain:	3-oxoacyl-[acyl-carrier-protein] synthase I homology <QAS>						
F534-815/Domain:	[acyl-carrier-protein] S-malonyltransferase homology <LADH>						
F1764-1944/Domain:	short-chain alcohol dehydrogenase homology						
F2015-2095/Domain:	acyl carrier protein homology <ACPI>						
<b>Query Match</b>							
Best Local Similarity	3.24;	Score 129.5;	DB 2;	Length 2110;			
Matches	151;	Conservative 82;	Mismatches 237;	Indels 289;	Gaps 33;		
Qy	151	EGIJDDEBATTVITDVF--WDK-----VPGALCVQIPEGGETEVILVRGEALLKRL	198	Db	29	BSLRGDK--LITEIPIPDRWDADDYDTPBPGFGR-SVSRNGG-----	68
Db	199	PSFKQDWVLAFLDVSDAMALDPPDFI--TCYEVFEALKLQLQERGASSIAPDLRAQI	255	Qy	69	--PLDDVAGFEEFGTISERHTSIDPOQRLILERSWEAIBHA-----	109
Qy	256	RAQIDETLEITPRYVTELLGIGPLGDDYAAKELINGLSGVRLTILWSVGGGASALVGGL	315				

Db 110 -----GLDPASL-----AGSSTAVFTGLTH 129  
 QY 316 BKKFNEBAFLMTRAEQNDLFVATPSNI-----PAESPEVY----- 350  
 Db 130 EDVI---VLTPTGGLASPVVTCGLNSVASGRIAHTLGLHGPMTDOPACSSGLMATH 186  
 QY 351 -----EVALALVAQAPIGKPHL-----ODADKOPQOLQAK 383  
 Db 187 ACPSLDHGERDLAAGCAVLLPHACVAASAQMLSSTRCRCSFDADAGFVSEGGAM 246  
 QY 384 VMMKETPAMYLYDTENNWEIDFGIQLCALLIGKVDECMMGLDSRDSQRYPRAIVEVY 443  
 Db 247 VLUKRLPDALDRGN---TFAVUVTGTTNQDET---LMPSEDQV---AVYRAA 295  
 QY 444 LEPNRRDDNDLPLGCKLUEBTWLAGVVFPRPRTDKKEPKLGQYDPPWLSLVEREVY 503  
 Db 295 LAAG-----VQPETVWV-----EHGTGPIGDPEYRALAR-YG 332  
 QY 504 OGSPLAAMATMARIQABHVKASAMALQKVP-----PSRYTDNSAEPKVOETVFS 555  
 Db 333 AGPICALGSKSAKSNNHGHTSASAGTVGLIKALSKSRHGVVPLPMLNRFLDELSV----- 386  
 QY 556 VDPVGNNGGRDGPGVFAEAVERSENFTNDIA-IRAGVSESVDECTVEMSVADMILK 614  
 Db 387 -----ETGLFVPUAVTPWPN-GNDHTPKRVAVSSFGMSGTVHATVEAPAE 432  
 QY 615 ASKILKLAQVA-GLISLPSQKFLKSSSSPQRKDMVSSMSREVATIGSVRADSEALPR 673  
 Db 433 ASPPESSEGGDAEY-----PRLMSSS-----SDALR 461  
 QY 674 MDARTAENTVSKMOK-----TKSLARGPDR-----TEMLPEVLGML 712  
 Db 462 QTRARLATWEEHRDCVAASDIALTARGERHRVRTAVAAVANPELVEGLRSVADGA 521  
 QY 713 -----KW-----TDAEAETPAQLG-----VYDTLILKLSVDSVTVSUDGTRALV 752  
 Db 522 YDAVGHDGDRGPWVFGCGSONHAMGTOALLASBPPVFAATIAKL-----EPVIAESEGFST 578  
 QY 753 EA-TLEBACLSLVLHPENNATVTRTYTRVETFWSKSG 790  
 Db 579 EATAQQTWTGKVKYQPAVFAVQVALATMEQTYGREG 617

RESULT 11

T49362 hypothetical protein B1D1.60 [imported] - *Neurospora crassa*  
 C;Species: *Neurospora crassa*  
 C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
 C;Accession: T49362  
 R;Schulte, U.; Aign, V.; Heineisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000  
 A;Reference number: Z25022  
 A;Accession: T49362  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1275 <SCH>  
 A;Cross-references: UNIPROT:Q9P6A4; EMBL:AL355927; GSPDB:GN00116; NCSP:B1D1.160  
 A;Experimental source: BAC clone B1D1; strain OR74A  
 C;Genetics:  
 A;Gene: NCSP:B1D1.160  
 A;Map position: 6  
 A;Intervall: 24/3

Query Match 3.2%; Score 129; DB 2; Length 1275;  
 Best Local Similarity 20.9%; Pred. No. 6.5; Mismatches 210; Indels 138; Gaps 26;  
 Matches 114; Conservative 84; MisMatches 114; Insertions 84; Deletions 114;

QY 325 RMTTAAEQLVPLVAPSNTIAPESFREVYEVALLVAQAPIGKPHL----- 368  
 Db 39 RNTTAPTPAYFALLSLNENLATPVVLLVTPF-APKPKIQAFTQITILAPVL 97

Db 98 LPDAD--APLMRASIGLE-SLILAQDPTQWELGTAQIGPRKAVAGLIA-----L 144  
 QY 426 GLSEDS-QYRNPAIVEFLENSNRDDMDPGLSCKLIEIWLAGVVFPRRTDKD----- 479  
 Db 145 ALDRPKYETKRADEBALNLINPPSPS----- 199  
 QY 480 -KKFKLGDYDDPMVLSYLERVEVQO---GSPLAAATM-----ARIGABEVKASAM 527  
 Db 200 RKEKKTIDSTHDPELILAHQLQVKAVASGGWFSKSKIESLCELLSIARTGNHNSMMAF 259  
 QY 528 QALQKVEFSSRYTR-NSAPKDVQETVSVDPYGN-----VGRDGEQGVFLAF 576  
 Db 260 EIFENMFEGMAADEVASAKLPLRPLRPAENDTQOLLPPWTAILSRAYD---VAGQ 315  
 QY 577 VRESENFE---TNDYAIRAGVSES-----SVBETVEMSVADMILKEASV-----KIL 620  
 Db 316 ISPAETFOLWDPELTVAGYLESOHKNIRVSSACLSLVSFLANCYPRKEAILEPSIFDEKVI 375  
 QY 621 AAGVAI--GLISLPSQKFLKSSSSPQRKDMVSSMSREVATIGSVRADSEALPR 669  
 Db 376 QOLVKVVEGULITWQOAHMETTFVULGAMDFARWOANPYLUSV--VNSIGERGNDF 432  
 QY 670 ALPMDARAAENTVSKMOK-----VJUGRMUWKWTDRAETAETQI 725  
 Db 433 A---GKOBADEVIGKATR---AMGPEAVLNVLPLNLAKPKVQGPGRW-----ML 476  
 QY 726 GLWVDT----LJKLSVDSVSVTSDGTRALVANTELESACLSLVLHPENNATVTRYTR 781  
 Db 477 PLRDYDENTNLHFKSILVPUASAIMFORVIE-----HOCEKNHNIKIEFIV 523  
 QY 782 YEVFTWS 787  
 Db 524 VQQIWS 529

RESULT 12

T33312 pilin biosynthetic Protein - *Pseudomonas aeruginosa*  
 C;Species: *Pseudomonas aeruginosa*  
 C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 04-Mar-2000  
 C;Accession: T30312  
 R;Ritchchurch, C.B.; Young, M.D.; Hobbs, M.; Mattick, J.S.  
 submitted to the EMBL Data Library, November 1996  
 A;Description: *Pseudomonas aeruginosa* chemotactic transduction genes pilI, chpA chpB and  
 A;Reference number: Z20819  
 A;Accession: T30312  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-836 <WHI>  
 A;Cross-references: EMBL:U79580; NID:93241967; PID:93241968; PIDN:AAC23930.1  
 C;Genetics:  
 A;Gene: pilI  
 C;Supergene: pilI  
 C;Supergene: *Pseudomonas aeruginosa* pilin biosynthetic protein  
 Query Match 3.2%; Score 128.5; DB 2; Length 836;  
 Best Local Similarity 20.0%; Pred. No. 3.6; Mismatches 158; Conservative 106; Gaps 35; Indels 235; Matches 158; Conservative 106; Mismatches 232; Indels 235; Gaps 35;

QY 124 DALISRROIQIAQCTESLNPSSRREYNEGLLDEEATWITDVPWMDKVGALCVLOEGCE- 182  
 Db 24 ETLXQAROLEAFTVENPODPTMR-----FCITYV--HQVQGTLOVNEFGCAA 69  
 QY 183 -TEIVURVGEALKERPKPSKPKDQVWMAFLDVSRRDAMALDPDF-----ITGYE 234  
 Db 70 LLAREMEDMVOALDGLPFGNQGRALVQALQPLDRIQTTARDLPLMVPLLAND 129  
 QY 235 FVERALKLQEGASSLAPD-----RAQD-----ETLEETPTVYELGLP 278  
 Db 130 RAARGEKUSE--TSLEFAPPDSLQRQDQGEAQLRDTDELGLRKURQDQMALVGL 187  
 QY 279 LGDDYAAK-----RNLGSGVRNT-LMSVGGGASALVGGLTREKFNMNEAFRLMT 327

Db	188 RNQDVATSGIYLARVARYARLEGICREAQPLWISI---ASGLYEGLANGSVVINSASVR-T	242	Qy	191 EAIIKLERLPSKSFQ--DVLVMALAFDLSRDA--MADLPDFITQYEFEALKULQE	245
Oy	328 AABQDV-----LFWTIPNSN-----PAEFEVVEVALVAQAFIG	363	Db	219 DEIKGERKLENKLRQRDLA LSKQLAACRCPAVELTD-DIVVKGDEBKVVALFQE	275
Db	243 LLRQDREKLRLVQEGADGGINQKAAPDELWKNLFLPYAKAPSOSPRIRAL-----	291	Qy	246 EGSSLAPLRAQIDE-----TELETPTVYLELIGI---PLGDDYAAKRLNG--	290
Oy	364 KKHLLQDA--DKQFQQLQOAKVNMAMEPAM-----LYDTTRNWEIDFGLERG--	409	Db	276 LGFOSFLDRMAVOTDEGEKPLAGMDFA TADSVDDEMIA DKAALVUVVGDNYHAP-VGI	335
Db	292 KQYRDLRPLDHPETDVADERARLLAGPDAMRSVVGALCCELVRIKOSLDLFRSDRGHP	351	Qy	291 -----LSGRVILMSVGGGASALVGGLTREKFMDAFLRMTAE---	330
Oy	410 --L-CALL--IGKODEBCRMWLGLEDSEDSQRNPALIVEFVLENSRDDND---DLPGLCKL	461	Db	336 ALANERGRFFLPETALADPKTLAW-----LGDETKCTMFDKRAVALWKGI	385
Db	352 SEDALLAPLKQIADTAVLGFSOPRKVILQDQDVLTHALQGRREPSAILMWDVAGALLY	411	Qy	331 -----QVDFLFVATPSNIPAS-----FEVIEVALVALVAQAFIGK-----KPHLQ	370
Oy	462 LETFLLAGVVFPP-RPRDTDKKKFLKGDDYD-----DPMLVSY-----	497	Db	386 ELRGVVFPLLAAYLLOPAQAGDVAVAKOMHQEAVRSDEAVYKGAKRTVDPDELA	445
Db	412 VELTAGMAGPGPERNSESRLPFTDVAQIHOQVKEARANGLEQAKDAIEFFIASQWNHE	471	Qy	371 DAQKQFOOLQOAKVNMAMEPAMALYDTRNNWE-IDFGLERGLCALLTG----KVDCRM	423
Oy	498 -----ERVENVOGS---PLAAMATMARIAGAHVKASAMQALQKVFP-----	535	Db	446 HLAR-----KAAAIWALEBPLMDBLRNNEQDRLLTEBQPLAGILANMFICVKVTKL	500
Db	472 HLARVPELITQVQGLAMIPLERBATTLETCNRYIQEQLL--ARKAVPDWQSHTLADAI	529	Qy	424 -WGLD-SEDSQYRNPAIVEFVLENSRDDNDLPGICKLJETMFLAGWVFERFDKDK	481
Oy	536 ---SRYTDNSAAPPKDVGETVFSY---DPVGN-----NVGRDGPBGVFIABVRPSENF	583	Db	501 BONGAELTEQO-----AVVEREYLQAGCQFBNINSP---KQJGT---VLFQKQLPVTK	549
Db	530 TSEYYELRLSERHASQSDLILVVAQDLSLANYGYTLKNSAEPGL-----	577	Qy	482 FKLGDDYDPPMLSYL-ERVENVOGSPLAAMATMARIAGAHVKASAMQALQKVFPRTD	540
Oy	584 ETNDYAIRAGVSSVDSNTVEMVSADMILKEASVYKILLAGVAGVALGLISLFSQKYFLKSSS	643	Db	550 TKIG-YSPSADVLEKLAPEHTE--HILHYRQLQKLOSTIE-GIKVWHRVTGKHTM	605
Db	578 -SCPAATESPAAPEPERPEAVE--VAETAOPPADTAPAEA-----	616	Qy	541 RNAEPKDQVQET--VFSVDPVCGNNVGRDGPERGVIFABVRSE-----NFETNDYAT	590
Oy	644 FQKDKMVSMSMESVATGSVRADDSEALPRMDARTAENIVSKWQKINSLAFGDPHRTML	703	Db	606 FNQR-----LTQGRLSSVEPNLQNIPIRLEGEGRKIKQAFVFPSEPDWLIFADYSQIELRV	661
Db	617 --REDAPOLASDQDNWLTGEVAPDAGE--PSLD-----ALIDLPLDDSAVP 658	Qy	591 RAGVSSVSDETVE-----MSYADMILE--ASVKAIAAGVAGVAGLUSLFSQ	634	
Oy	704 PEVLDGRMLKIKWDRAAETAQOLGLVYDYLKLISVDSVTSVAD-----GTRALVEATL	758	Db	662 LAHIAE-----DDMIAEATRGRGDIQHTKAMDIFHVHSBDVTANMRKOKAANGVIGI	718
Db	659 PAVPECRE---TPPQSTSAPASLDDFSLDEIDLSDLPAADAAPASGPAAALADSLPE	714	Qy	635 KYFLKSSSSFORQDMVSSMSMESVATGSVRADDSEALPRMDARTAENIVSKWQKIKSLF	694
Oy	759 SACL-SDVLHP 768	715 QMGLGDDBLAQP 725	Db	719 -YGLAQUNLITRKEAAEFIERFASPEPGVK-----QYMNNTIODE-AKQGYV	764
Db	715 QMGLGDDBLAQP 725		Qy	695 GPDHRTMLEPEVLDGRM-LKWTIDRA-ETMQOLGLVYDYLKLISVDSVTSVADGTA-	750
RESULT 13			Db	765 TLIHRRYLPDITSRNRVNRSPARTMTPIQGSAD-IKKAMIDLSVRLREERLQAR	823
S70368	DNA polymerase I - <i>Bacillus stearothermophilus</i>		Qy	751 -----LVEATLRESACASDLYVPE	769
C;Species:	<i>Bacillus stearothermophilus</i>		Db	824 LLIQVHDBLILIEAPKSERIERICLRLV-PE	850
C;Accession:	S70368				
C;Accession:	S70368				
A;Status:	preliminary				
A;Molecule type:	DNA				
R;Riggs, M.G., Tudor, S., Sivaram, M., McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368; MUID:9628813; PMID:8679703				
C;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I				
A;Accession:	S70368				
R;Riggs, M.G., Tudor, S., Sivaram, M.; McDonough, S.H.					
R;Blochim. Biophys. Acta 1307, 178-186, 1996					
A;Title:	Construction of single amino acid substitution mutants of cloned <i>Bacillus stearothermophilus</i> DNA polymerase I		</		

Matches 159; Conservative 105; Mismatches 291; Indels 239; Gaps 36;

Qy 124 DALISRROITQACETLSNRSRISRSRBYNEGILIDDEATVITDVPNDKVPGALCIVQEGSR- 182  
 Db 19 ETIKQARQALEAVENPQDPTMR-----FCLTYV--HQVQSTLQWVFYGA 64

Qy 183 --TEIVLURGEALIJKERLIPSKFKQDVVLMALAPLVDSDAMALDPDF----ITGYE 234  
 Db 65 LLABEMEOLVOAIIJLDRGVERNOGEALEVLMQALOLPVYLDRIOPARROLPMVNUPLNL 124

Qy 235 FVEEALKLIQOBEGASSLADL--RAQID-----ETLEBETTPRYVUELLGP 278  
 Db 125 RAARGEKLISE--TSLFADDLSQRPQDGEATIQQLRTBEGJLJKURQTOQMAVGIL 132

Qy 279 LGDDYAAK-----RUNGSCVRNT--LWSVGGGASALVGGLTRKFMBAFLRM 327  
 Db 183 RNQDVATSLGYLARVARLEGLCRAPLGLPLWSI---ASGLVGLANGSVWVISASVR-T 237

Qy 328 AAEQVD-----LFVAMPNT-----PAESPEVYELAVLAQAFIG 363  
 Db 238 LLRQDRELKLRLVBOGADGLENQAJAQPDELVNLFLFYVAKAPSPIRAL----- 286

Qy 364 KKEPHILQDA--DKOFQQLQAKVNAMEI-----LDTTRNWEIDGELRG-- 409  
 Db 287 KEQYTRIDEALPDHETVDAARRLAGPDRAMRWSVAGLCELRIRKIDSIDFLFRSDRHP 346

Qy 410 --LCAALL--IGKQDRCRMWLGLOSDSQYRNPAVEFVLENSHRDD---DLPGLCKL 461  
 Db 347 SELDALLAQLQIAADTLAVLGFGQPKVILQDQDVTIALAQGRREPSDAILMDVAGALLY 406

Qy 462 LETWLLAGVVF-PFRDTKDKKKFKGDYD-----DPMLSYJ----- 497  
 Db 407 VEATLAGMAGPGDRNBSRSLPITDVAQIHOQVTEKARINGLFLQAKDATERIASQWNHE 466

Qy 498 -----ERVEVGGS-----PLAAMATMARIAGAENVKASAMQALQVFP----- 535  
 Db 467 HILARVPEBLTQVCGGLAMPLERATLTCTCNYIQEQIL--ARKAVPDWQSIDLADAI 524

Qy 536 ---SRYTPRNSAEPKDVOFTVFVSY--DPVGN-----NVGRDCBPGYFTAEAVRPSNF 583  
 Db 525 TSVYELYLSELSEPHASQSDLILIDVAEDSLANGLYTLENKPNSSAEPGL----- 572

Qy 584 ETNDYATRAGVSESSVDETTEVNSVADMILKEASVTKILAGVAGLISFSQVKPLKSSS 643  
 Db 573 :-----SPGATEESPAAPEPERPEAVE--VAETAQOPADATPAREA----- 611

Qy 644 FORKDMVSSMSDVATGVSRAVDSEALPRMDARTAENIVSKWOKIKSLAFGDHREM- 702  
 Db 612 --KRDAPQQLASDOWNTLGVAAPAGE--PSLIL-----ALDPLPDSALV 653

Qy 703 --LPEVLQDRMLKWTDRDAAETAOGLGVYDVTIILKLSVSVTSAD----GTRALVET 755  
 Db 654 PALPEVVVERSGQPOSTSAPARSID----DPSIDRIDIISGLDIPADAAPASGPALADWS 708

Qy 756 LEBSACT-SPLVNP 768  
 Db 709 LPQWGLGDDLAQ 722

A;Cross-references: UNIPROT:Q9A779; EMBL:AB003696; NID:di106e37; PID:di1020958; PIDN:BA2C

Query Match 3.1%; Score 126.5; DB 2; Length 1018;  
 Best Local Similarity 19.2%; Pred. No. 6.6;  
 Matches 128; Conservative 109; Mismatches 265; Indels 163; Gaps 28;

Qy 115 KPPRGPSDALLSRRQITQACETLSNRSRISRSRBYNEGILIDDEATVITDVPNDKVPGAL 174  
 Db 374 EPPQKQQVPPITPQNQLP--ETSQMPQS-QQVNPNLIN--AATALGSMODLINTVN 428

Qy 175 CVLOBGETTEBIVLAVGEALKERLIPSKFKQDVVLMMA--LAFIDVSRDAMALDPDFITG 232  
 Db 422 AGLK--EDLILKEATAALNDK----KSDIAEKQANTIALENTVUNKNLUPDAKG 481

Qy 233 YEFVEEALKLIQOBEGASSLADP-----AQIDTBLEBITPRVYLEL--LGIP 278  
 Db 482 VNAVLETINDON-----TPDLIEKSQLEAVTAIALNSBNLSPKQKOCILEKAVDVGIL 535

Qy 279 LGDDYAAKTLNGLGSVGRNTILWSVGGGASALVGGLTRKFMBAFLR--MTAEEQDVFV 336  
 Db 536 LKOD-----ASRAAIDGIT-----DAVIKSNLSTDKGTMF 569

Qy 337 ATPSNIPARSPFEVETALLVQAOFQKPHLQDQKFOOQOKUMA-----ME 388  
 Db 569 AvgdKUNVSELSNAEKQKLGSVLKKGTVQASPAQQLMQNLDKTAEQTKKDTK 628

Qy 389 IPAMLYDTUNNWE-----IDFGLERGICALLIGKVDBCRMWLGJLQEDSQR 435  
 Db 629 VNDILFDPLSNTELKTWTQAIISNVLDGCPATEVKGBIIOETNTVAGSSLRHD--- 684

Qy 435 NPATVFEVLEN--SNRDNDPQGCKLQETLCTQLAGVFFRFRUTDKKFKLGYYDDMV 493  
 Db 685 KAATIKGIBETIATHHSDSLISLNKALINMASAKGIA-----ESQNLNPDR-RELMT 734

Qy 494 LSYLERVEVQGSPLAAMATMARIAGAENVKASAMQALQ--VFPSPRTDRNSA----- 544  
 Db 735 KGLUDGIVEGKGKGPFRPEITKAVKUSSIDNSNINDSKKEALKKAKDASEALDRGTQNTJEGL 794

Qy 545 -----EPKD-----VQVTVSVDPVGNNVGRDGERGVFIAE----- 575  
 Db 795 KGONIEEHKPHDDIYINKYKVRVINAUNPVTEALEKSKPSVVSABERIVOTESTLANNISKL 854

Qy 576 AVPSENTEFTNDYAIRAGVSESSVDETTEVNSVADMILKEASVTKILAGVAGLISFSQ 635  
 Db 855 AVEKVNPF-----RAMLSPNGNLKTLEK-----KEESIKKVE-----LVKARTK 896

Qy 636 YFLKSSSSFRQKDMVS----SMESDVATGVSRAADD--SEALPRMDARTAB-NIVSKWQK 688

Qy 897 SSTRQQQSFFIKANLIDDKTLKSLKIRQLTDIKLUQEQRGAETENPSVKTEDVRWSKSK 956

Qy 689 IKSLA 693

Db 957 LKPIS 961.

RESULT 15

T30853 antigenic heat-stable protein - *Rickettsia japonica*

C;Species: *Rickettsia japonica*  
 C;Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 09-Jul-2004

C;Accession: T30853  
 R.Uchiyama, T.  
 submitted to the EMBL Data Library, May 1997  
 A;Description: sequencing of the gene encoding the antigenic heat-stable 120-kilodalton

A;Reference number: 220905  
 A;Accession: T3053

A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Residues: 1-1018 <UCH>

Search completed: June 10, 2005, 01:42:11  
 Job time : 34 secs

THIS PAGE BLANK (USPTO)



QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420	Db	121 FSDALISRQIQAACTLSNPSRRBYNSGCLLDDGATVITDVPMWDKVPGALCVI QEG 180
Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420	QY	181 GETEIVLRVGEALKERLKPKSQRQDVTLVMAALFDYDSRANALDPDFPRCYEFEAL 240
QY	421 CRMWLGJEDSEDSQYRNPAIVEFVLENSNRDDNDLPGCKLLETWLAGVVFPRFRDTKD 480	Db	181 GETEIVLRVGEALKERLKPKSQRQDVTLVMAALFDYDSRANALDPDFPRCYEFEAL 240
Db	421 CRMWLGJEDSEDSQYRNPAIVEFVLENSNRDDNDLPGCKLLETWLAGVVFPRFRDTKD 480	QY	241 KLIQEGBASLADPDIRQIDEETTPRYVLELGLPLGDDYAARLNGLSGVRNLWS 300
QY	481 KFKLGDYDDPMVLSYLVERVEVQGSPLAATMARIAGAHLHVAKASAMOALQKVFPSPRTD 540	Db	241 KLIQEGBASLADPDIRQIDEETTPRYVLELGLPLGDDYAARLNGLSGVRNLWS 300
Db	481 KFKLGDYDDPMVLSYLVERVEVQGSPLAATMARIAGHVAKASAMOALQKVFPSPRTD 540	QY	301 VOGGASBLVGLTRECEMNEFLRMTAAEQDLFVATPSNIPAESPEVYEVALVAQA 360
QY	541 RNSAEPKVOETVFSVDPGVNNYGRDGERGPFVAFTRPSENFTNDYAIRAGVSESSVD 600	Db	301 VEGGGASALVGLLTRECEMNEFLRMTAAEQDLFVATPSNIPAESPEVYEVALVAQA 360
Db	541 RNSAEPKVOETVFSVDPGVNNYGRDGERGPFVAFTRPSENFTNDYAIRAGVSESSVD 600	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
QY	601 ETIVEMSVADMILEKEASVYLQKLAQGAIGLISLFSQYKLUKSFFORKDMVMSMESDATI 660	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
Db	601 ETIVEMSVADMILEKEASVYLQKLAQGAIGLISLFSQYKLUKSFFORKDMVMSMESDATI 660	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
QY	661 GSVRADSEALPRMDARTAENTIVSKWQKIKSLAFIGDPDRIMPEVLDGRMLKTIWDR 720	Db	421 CRMWLGJEDSEDSQYRNPAIVEFVLENSNRDDNDLPGCKLLETWLAGVVFPRFRDTKD 480
Db	661 GSVRADSEALPRMDARTAENTIVSKWQKIKSLAFIGDPDRIMPEVLDGRMLKTIWDR 720	QY	481 KFKLGDYDDPMVLSYLVERVEVQGSPLAATMARIAGAHLHVAKASAMOALQKVFPSPRTD 540
QY	721 ETAQGLGYDVTYLLKLSVTSVADGTRALVATEESACSLDVLVPPENNATDVRYT 780	Db	481 KFKLGDYDDPMVLSYLVERVEVQGSPLAATMARIAGAHLHVAKASAMOALQKVFPSPRTD 540
Db	721 ETAQGLGYDVTYLLKLSVTSVADGTRALVATEESACSLDVLVPPENNATDVRYT 780	QY	541 RNSAEPKVOETVFSVDPGVNNYGRDGERGPFVAFTRPSENFTNDYAIRAGVSESSVD 600
QY	781 RVEVFWMSKGWKITEGSVLS 801	Db	541 RNSAEPKVOETVFSVDPGVNNYGRDGERGPFVAFTRPSENFTNDYAIRAGVSESSVD 600
Db	781 RVEVFWMSKGWKITEGSVLS 801	QY	601 ETIVEMSVADMILEKEASVYLQKLAQGAIGLISLFSQYKLUKSFFORKDMVMSMESDATI 660
<b>RESULT 2</b>			
ID	OTXAR9 PRELIMINARY; PRT; 801 AA.	Db	601 ETIVEMSVADMILEKEASVYLQKLAQGAIGLISLFSQYKLUKSFFORKDMVMSMESDATI 660
AC	OTXAR9;	QY	661 GSVRADSEALPRMDARTAENTIVSKWQKIKSLAFIGDPDRIMPEVLDGRMLKTIWDR 720
DT	01-OCT-2003 (TREMBlre. 25, Created)	Db	661 GSVRADSEALPRMDARTAENTIVSKWQKIKSLAFIGDPDRIMPEVLDGRMLKTIWDR 720
DT	01-OCT-2003 (TREMBlre. 25, Last sequence update)	QY	721 ETAQGLGYDVTYLLKLSVTSVADGTRALVATEESACSLDVLVPPENNATDVRYT 780.
DT	01-MAR-2004 (TREMBlre. 26, Last annotation update)	Db	721 ETAQGLGYDVTYLLKLSVTSVADGTRALVATEESACSLDVLVPPENNATDVRYT 780
DE	division protein.	QY	781 RVEVFWMSKGWKITEGSVLS 801
GN	Name=ARC6;	Db	781 RVEVFWMSKGWKITEGSVLS 801
OS	Arabidopsis thaliana (Mouse-ear cress).	QY	781 RVEVFWMSKGWKITEGSVLS 801
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	Db	781 RVEVFWMSKGWKITEGSVLS 801
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	QY	781 RVEVFWMSKGWKITEGSVLS 801
OC	eurosid II; Brassicales; Brassaceae; Arabidopsis.	Db	781 RVEVFWMSKGWKITEGSVLS 801
OX	NCB_ TAXID=3702;	QY	781 RVEVFWMSKGWKITEGSVLS 801
RN	[1]	Db	781 RVEVFWMSKGWKITEGSVLS 801
<b>RESULT 3</b>			
ID	OFIG9 PRELIMINARY; PRT; 801 AA.	Db	121 FSDALISRQIQAACTLSNPSRRBYNSGCLLDDGATVITDVPMWDKVPGALCVI QEG 180
AC	OFIG9;	QY	181 GETEIVLRVGEALKERLKPKSQRQDVTLVMAALFDYDSRANALDPDFPRCYEFEAL 240
DT	01-MAR-2001 (TREMBlre. 16, Created)	Db	181 GETEIVLRVGEALKERLKPKSQRQDVTLVMAALFDYDSRANALDPDFPRCYEFEAL 240
DT	01-MAR-2001 (TREMBlre. 16, Last sequence update)	QY	241 KLIQEGBASLADPDIRQIDEETTPRYVLELGLPLGDDYAARLNGLSGVRNLWS 300
DT	25-OCT-2004 (TREMBlre. 28, Last annotation update)	Db	241 KLIQEGBASLADPDIRQIDEETTPRYVLELGLPLGDDYAARLNGLSGVRNLWS 300
DE	Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MDH9	QY	301 VOGGASBLVGLTRECEMNEFLRMTAAEQDLFVATPSNIPAESPEVYEVALVAQA 360
GN	Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MDH9	Db	301 VEGGGASALVGLLTRECEMNEFLRMTAAEQDLFVATPSNIPAESPEVYEVALVAQA 360
DB	(Hypothetical protein Atsg42480); Name=Atsg42480;	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
DR	EMBL: AY21469; AAQ16646_1; -	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
DR	InterPro; IPR0061623; Dnaj N.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
SO	SEQUENCE 801 AA; 88247 MW; T2E1B3FD4BE4B61 CRC64;	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
<b>Query Match</b>			
Best Local Similarity	99.8%; Score 4052; DB 2; Length 801;	Db	121 FSDALISRQIQAACTLSNPSRRBYNSGCLLDDGATVITDVPMWDKVPGALCVI QEG 180
Matches	799; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	QY	181 GETEIVLRVGEALKERLKPKSQRQDVTLVMAALFDYDSRANALDPDFPRCYEFEAL 240
QY	1 MEILISHVIGLSPOLCRUPPATKLKRSHNTSTICSASKWADRLLSDFNPSDSSSS 60	Db	181 GETEIVLRVGEALKERLKPKSQRQDVTLVMAALFDYDSRANALDPDFPRCYEFEAL 240
Db	1 MEILISHVIGLSPOLCRUPPATKLKRSHNTSTICSASKWADRLLSDFNPSDSSSS 60	QY	241 KLIQEGBASLADPDIRQIDEETTPRYVLELGLPLGDDYAARLNGLSGVRNLWS 300
QY	61 RATATTATLVSPPSIRPERPVPIDPYQVIGAQTHFLGIRRAFEARVKPQFG 120	Db	241 KLIQEGBASLADPDIRQIDEETTPRYVLELGLPLGDDYAARLNGLSGVRNLWS 300
QY	61 RATATTATLVSPPSIRPERPVPIDPYQVIGAQTHFLGIRRAFEARVKPQFG 120	Db	301 VOGGASBLVGLTRECEMNEFLRMTAAEQDLFVATPSNIPAESPEVYEVALVAQA 360
Db	121 FSDALISRQIQAACTLSNPSRRBYNSGCLLDDGATVITDVPMWDKVPGALCVI QEG 180	QY	301 VEGGGASALVGLLTRECEMNEFLRMTAAEQDLFVATPSNIPAESPEVYEVALVAQA 360
RP	[1]	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	QY	361 FICKKPHILQDAKQFOOLQOAKUMAMEI PAMYDTRNNWEIDFGLERGLCALIGKDE 420
RP	SEQUENCE FROM N.A.	Db	361 FICKKPHILQ

RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,  
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,  
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Meyers M.C., Marland M., Narusawa M., Nguven M., Palm C.J.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,  
 RA Davis R.W., Becker J.R., Theologis A.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M.,  
 RA Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,  
 RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusawa M.,  
 RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,  
 RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Becker J.R.,  
 RA Theologis A.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB016888; BAB1089.1; --.  
 DR EMBL; AV01395.1; --.  
 DR EMBL; AX15030; ANI1207.1; --.  
 DR InterPro; IPR001623; DnaJ\_N.  
 KW Hypothetical protein.  
 SQ SEQUENCE 801 AA; 88259 MW; 608E776FBA73PBCF CRC64;  
 Query Match 99.7%; Score 4051; DB 2; Length 801;  
 Best Local Similarity 99.8%; Pred. No. 3.8e-240;  
 Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 MEALSHVIGLSPOLCRLPPATKLRSHTNSTTICASKWADRLLSDPNFTSDSSSS 60  
 1 MRALSHVGIGLSFOLCLRLPPATKLRSHTNSTTICASKWADRLLSDPNFTSDSSSS 60  
 OY 61 FATTATTTVLSPSISDPRPERAVIPIDFYQVLAQTHFLTGIRRAFEARVKPPORG 120  
 61 FATTATTTVLSPPSISDPRPERAVIPIDFYQVLAQTHFLTGIRRAFEARVKPPORG 120  
 DB 181 GEERIVLVRGEALKERIKERLPSKSPQDVULMALDVSRA DAMALDPDPRTGEPVEAL 240  
 181 GEERIVLVRGEALKERIKERLPSKSPQDVULMALDVSRA DAMALDPDPRTGEPVERAL 240  
 DB 121 PSDDALTSRROIQIQAETLNSNRSRENEYNEGLIDDEBEATVTIDPVDPKVGA CVLOEG 180  
 121 PSDDALTSRROIQIQAETLNSNRSRENEYNEGLIDDEBEATVTIDPVDPKVGA CVLOEG 180  
 OY 241 KLIQBEGASSLADPLRAQIDETEITPRYVLELGLPLGDDYAKRNLGSVNTIWS 300  
 241 KLIQBEGASSLADPLRAQIDETEITPRYVLELGLPLGDDYAKRNLGSVNTIWS 300  
 DB 301 VGGGGASALVGGITREKFENEARMTAQBQVLFATPSNITARESEFTEVALLVAQ 360  
 301 VGGGGASALVGGITREKFENEARMTAQBQVLFATPSNITARESEFTEVALLVAQ 360  
 OY 361 PIKSKPHILQDAKQFOOLQAKUNAMEPAMYDTRNNWRIGLERCICALIGKDE 420  
 361 PIKSKPHILQDAKQFOOLQAKUNAMEPAMYDTRNNWRIGLERCICALIGKDE 420  
 DB 421 CRMWGLDOSBDSQRNPATVEFLIENSRRDDPLGLICKLSTWLAGVFPFRDTDK 480  
 421 CRMWGLDOSBDSQRNPATVEFLIENSRRDDPLGLICKLSTWLAGVFPFRDTDK 480  
 DB 421 CRMWGLDOSBDSQRNPATVEFLIENSRRDDPLGLICKLSTWLAGVFPFRDTDK 480  
 OY 481 KPKLGDYDPMWLSYLRSVEVYQGSPPLAAATMARGHEHVKSAMOLOKTPPSRTD 540  
 481 KPKLGDYDPMWLSYLRSVEVYQGSPPLAAATMARGHEHVKSAMOLOKTPPSRTD 540  
 OY 541 RNKRPKVQVOETTSVDPVGNNGRDRGECGVPAEAVPSENFTNDYAIRAGVSESDV 600  
 541 RNKRPKVQVOETTSVDPVGNNGRDRGECGVPAEAVPSENFTNDYAIRAGVSESDV 600  
 DB 601 ETIVEMSVADMILKEASVKLAA3VAGLTSLSFSOKVFKSSSFQRKOMVSSMSDVTI 660  
 601 ETIVEMSVADMILKEASVKLAA3VAGLTSLSFSOKVFKSSSFQRKOMVSSMSDVTI 660  
 DB 601 ETIVEMSVADMILKEASVKLAA3VAGLTSLSFSOKVFKSSSFQRKOMVSSMSDVTI 660  
 DB 601 ETIVEMSVADMILKEASVKLAA3VAGLTSLSFSOKVFKSSSFQRKOMVSSMSDVTI 660  
 OY 661 GSYRADDSBALPRMDARTAENTVSKWQKITSLSAFGPDIREMPEVLDGRMLKIWTDRA 720  
 DB 661 GSYRADDSBALPRMDARTAENTVSKWQKITSLSAFGPDIREMPEVLDGRMLKIWTDRA 720  
 OY 721 ETQOLGIVYDYLKLSVSUTVSADGTRALVETLRSACSLVIRENNATDVRTYT 780  
 DB 721 ETQOLGIVYDYLKLSVSUTVSADGTRALVETLRSACSLVIRENNATDVRTYT 780  
 OY 781 KIEVFMWSKGKWTGESVLS 801  
 DB 781 KIEVFMWSKGKWTGESVLS 801  
 OY 801 RYEFVWSKGKWTGESVLS 801  
 DB 801 RYEFVWSKGKWTGESVLS 801  
 RESULT 4  
 ID 062729 PRELIMINARY; PRT; 760 AA.  
 AC 062729 DT 05-JUL-2004 (TREMBREL 27, Created)  
 AC 062729 DT 05-JUL-2004 (TREMBREL 27, Last sequence update)  
 DE PLASTID DIVISION PROTEIN.  
 DR Name=PO575F10.2;  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Bokaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Sporophytina; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Saeki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(Ga3) genomic DNA, chromosome 2, PAC  
 clone:PO575F10.2";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB004885; BAD07642.1; --.  
 DR InterPro; IPR001623; DnaJ\_N;  
 SQ SEQUENCE 760 AA; 84134 MW; 2C4684862795B2F CRC64;  
 Query Match 43.7%; Score 1775.5; DB 2; Length 760;  
 Best Local Similarity 47.9%; Pred. No. 2.1e-100;  
 Matches 390; Conservative 119; Mismatches 213; Indels 93; Gaps 16;  
 OY 12 SPPOLCRPATKLRSHTNSTTIC-SASKWADRLLSDPNF---TSDSSSSFAT 65  
 14 APFAFSLPRPRPRPRPRPRPHPSAACRASWRFLADPHLPTAAPSDPPSPAPAPA 73  
 DB 66 TTATLSPSISDPRPERAVIPIDFYQVLAQTHFLTGIRRAFEARVKPPORG 125  
 74 APSASPFELPDAERSLPLQDFYKULGABRHPLGIDGIRRAFEARAKPQGYGSDA 133  
 OY 126 LSRRQIQAETLNSNRSRENEYNEGLIDDEBEATVTIDPVDPKVGA CVLOEGGET 185  
 134 LVGRQMLQIAHDTLMQNRSQTODRALSENREBALTWDIANDK-----EAGEALA 184  
 DB 186 LVRGEALKERIKERLPSKSPQDVULMALDVSRA DAMALDPDPRTGEPVEALKLG 245  
 188 LVLTGEOQLIDRPPRKFQDVULMALVYUDLSRDAMASPPDVIGCEVLERAKLQ 244  
 DB 246 EGSSLADPLRAQIDETEITPRYVLELGLPLGDDYAKRNLGSVNTIWSVGGG 305  
 245 DGAISNLADDILSISDTEITPRYVLELGLPLGDDYAKRNLGSVNTIWSVGGG 304  
 DB 306 ASALVGGSLTREKFENEARMTAQBQVLFATPSNITARESEFTEVALLVAQ 365  
 DB 305 IATVGGGSRFAFMNEFLRMITSQEDFFSKTPNSIPPEWFBFVYVALVAQAIISK 364  
 OY 365 PHLOQAKQFOOLQAKUNAMEPAMYDTRNNWRIGLERCICALIGKDECRMW 425  
 DB 365 POFIMMADLDRPQKENGNS---HYAVDN---EMDLALERAFCSLIVGDVKCRMW 416  
 OY 426 GLDSBDQSQRNPATVEFLIENSRRDDPLGLICKLSTWLAGVFPFRDTDK 484  
 DB 417 GIDNESSPFRDPKILEFTVNNSISEENDLPLCKLJETWLJFVEVPRSRDTRGMQFL 476



QY	1 MEALISHVGIGLSPRQLCPRLPATTKRRSHNTSTTCASKWADRLSLPNFTSDSSSS 60	Db
QY	1 MEALISHVGIGLSPRQLCPRLPATTKRRSHNTSTTCASKWADRLSLPNFTSDSSSS 60	Db
QY	61 FATTATTLVSLPPSIDIQRPERHYPIDPYQVIGAQTHFLDCCRRAFEARVEKPPORG 120	Db
QY	61 FATTATTLVSLPPSIDIQRPERHYPIDPYQVIGAQTHFLDCCRRAFEARVEKPPORG 120	Db
QY	121 FSDDALSRQILOACETLSNPRSRREYNEGLDDEATWITDPWDKVPGALCYLQBG 180	Db
QY	121 FSDDALSRQILOACETLSNPRSRREYNEGLDDEATWITDPWDKVPGALCYLQBG 180	Db
QY	181 GETEIVLRVGEALLEKLERIKPSFKODVVLVALAFDLSRDAMALDPPDRITGYFVEREAL 240	Db
QY	241 KLIQESEGSSLAPDILRAQIDETAREITPRYVLELGLPLGDDYAKRKLNGLSGRNILMS 300	Db
QY	241 KLIQESEGSSLAPDILRAQIDETAREITPRYVLELGLPLGDDYAKRKLNGLSGRNILMS 300	Db
QY	301 VGGGGASALVGGLTREKPNEAFL 324	Db
QY	301 VGGGGASALVGGLTREKPNEAFL 324	Db
RESULT 7		
OBYTLO	PRELIMINARY; PRT; 798 AA.	
ID	OBYTLO;	
AC	DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)	
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)	
DE	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	
DE	012707 protein.	
OrderredLocusNames=al12707;		
OS	Anabaena sp. (strain PCC 7120).	
OC	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.	
OX	NCBI_TAXID=103690;	
RN	[1] -	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=21595285; PubMed=11759840;	
RX	Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,	
RX	Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,	
RX	Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M., Yasuda M., Tabata S.;	
RX	"Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Anabaena sp. strain PCC 7120.";	
RL	DNA Res. 8:205-213(2001); EMBL; AP003590;	
DR	PIR; AD2144; AD2144.	
DR	HSSP; P08622; 1B0Z.	
DR	Pfam; PR00226; Dnaj; 1.	
DR	SMART; SM00271; Dnaj; 1.	
DR	PROSITE; PS50076; DNAJ_2; 1.	
KW	Complete proteome.	
SQ	SEQUENCE - 798 AA; 90124 MW; 3C57E6C221/722AA CRC64;	
Query Match	12.2%; Score 497; DB 2; Length 798;	
Best Local Similarity	24.6%; Pred. No. 7, 96-22;	
Matches	207; Conservative 139; Mismatches 301; Indels 196; Gaps 33;	
QY	84 VPTPIDFVQVLGAQTHFLDGDGIRRAFEAVSKPKPQQFGFSDDALSRQILOACETLSNP 143	Db
QY	11 VRIPLDYTRILGSLPLAASDEQKRYNSDRIVQPLPREYSQASIAKSRQKOLIEAVTVLSDP 70	Db
QY	144 RERREY-----NEGLDDEATWITDPWDKVPGALCYLQBG 176	Db
QY	71 KERSSYDQYLAHAYDPDNAAKKWAVENRGDSNKGHFVQSI--EVSEELIGALLI 128	Db
QY	177 LQEGGETIVLRVGEALL-----KERIKPSF---KQDVVLVALDFDSR 219	Db
QY	129 LQBIGBETIVLRVGEALL-----KERIKPSF---KQDVVLVALDFDSR 219	Db
RESULT 8		
O7NFP3	PRELIMINARY; PRT; 626 AA.	
ID	O7NFP3	
AC	DT 01-MAR-2004 (TREMBLrel. 26, Created)	
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)	
DE	01-MAR-2004 (TREMBLrel. 26, last annotation update)	
DE	Glr481 protein.	
GN	OrderredLocusNames=glr3481;	
OS	Gloebacter violaceus.	
OC	Bacteria; Cyanobacteria; Chroococcales; Gloebacter.	
OX	NCBI_TAXID=33072;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=PC_7421;	
RC	MEDLINE=2297700; PubMed=14621292;	
RA	Nakamura Y., Kaneke T., Sato S., Mimuro M., Miyashita H., Tsuchiya T., Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S., Takeuchi C., Yamada M., Tabata S.;	
RT	"Complete genome structure of the Gloeobacter violaceus PCC 7421, a cyanobacterium that lacks thylakoids.";	
RL	DNA Res. 10:137-145(2003); EMBL; AP00650; Rar01422; 1.	

DR InterPro; IPR001623; Dnaj\_N.  
 DR InterPro; IPR00841; IPR-Tike.  
 KW Complete proteome.  
 SQ SEQUENCE 626 AA; 69297 MW; 372BABB2588B4629 CRC64;  
 Query Match 11.0%; Score 448.5; DB 2; Length 626;  
 Best Local Similarity 25.8%; Pred. No. 5.3e-19; RT [1]  
 Matches 192; Conservative 112; Mismatches 282; Indels 159; Gaps 30;  
 86 IPIDFYQVLGAQTHFLTDGIRRAFEARSKPKPQFGSPDSLISRQLQACETLSPRS 145  
 9 LPSYVQIGVPPQCTYEQPERAFDRIAQAPREFSAVRSKPKPQFGSPDSLISRQLQACETLSPRS 145  
 QY RREYN-EG---LIDDEBEATVTIDVPMWDKVPGALCVTQEGGETEIVRVEAHLKILRK 200  
 146 RREYN-EG---LIDDEBEATVTIDVPMWDKVPGALCVTQEGGETEIVRVEAHLKILRK 200  
 QY |:|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 69 REIYHREGKQGLLIDSSHRAV----GLFLYLYELGEVOTLIERQREALAVADQP- 117  
 Db RA Koksharova O.A., Wolk C.P.;  
 DR "A novel gene that bears a Dnaj motif influences cyanobacterial cell  
 SEQUENCE FROM N.A. division."  
 RT J. Bacteriol. 184:5524-5528 (2002).  
 RX DOI=10.1128/JB.184.19.5524-5528.2002;  
 RX MEDLINE=200206501; PubMed=12218043;  
 RX DR InterPro; IPR001623; Dnaj\_N.  
 RP DR pfam; PF00226; Dnaj; 1.  
 DR EMBL; AR42196; RALI6071; SMART; SM0071; PROSITE; R50076; Dnaj\_2; 1.  
 DR PROSITE; R50076; Dnaj\_2; 1.  
 DR Cell division; KW DR pfam; GO:000910; P:cytokinesis; IEA.  
 DR PROSITE; R50076; Dnaj\_2; 1.  
 DR SEQUENCE 631 AA; 70693 MW; 6B46142E3B48DEB2 CRC64;  
 SQ OX NCBI\_TaxID=1140;  
 RN [1]  
 SEQUENCE FROM N.A.  
 Query Match 10.9%; Score 442; DB 2; Length 631;  
 Best Local Similarity 24.5%; Pred. No. 1.3e-18; RT [1]  
 Matches 189; Conservative 118; Mismatches 262; Indels 204; Gaps 30;  
 86 IPIDFYQVLGAQTHFLTDGIRRAFEARSKPKPQFGSPDSLISRQLQACETLSPRS 145  
 3 IPIDFYQVLGAQTHFLTDGIRRAFEARSKPKPQFGSPDSLISRQLQACETLSPRS 145  
 QY |:|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 169 TIKRWRPERILQILQA-GAAADPSPQRQEGMALLAALIAERSEGIEGDCNDQ--SGSRE 225  
 261 ETIEEITRTRYVLELLGLPLGDDYAKRKLGSVRNL--WVGEGGASALVGLTRE- 316  
 317 --KEFMNEAFLRMVTAEQVQLDF--VATPSNIPAESFBEVYEAVALVAQAFIGKPHILQD 371  
 QY |:|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 226 FVQFLQVLRLLTVQELFERARRS--PAQ--YLAQQLACGFTEGSPQCRR 280  
 372 ADQFOQOLQOAKVMAMPIALYDTRNNWEIDGLERLCALIGKVDCRMWLGQED 431  
 QY |:|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 281 ARCHLKLKVORO-----DNNEELAVCALLIGQVEAOKNERSABE 321  
 432 SQYRNPAIVEFLENNSRDDNDLPGICKLUEWLAGVPPRERDTKDKFKLGDYDDP 491  
 QY |:|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 322 Q-----AVDYL-KNLSDSPDPLPGICRYTDLWLABEVPGFRDRRSRTKLYFAHP 374  
 Db 195 KERLKPSFKQDVTLVMAFLDVSROMALAPPDRITGYETVEALKLLOQEGA---- 248  
 Db 3 OSALADHORFEPSPSLSRTHQEDQPLRPTILRLAQLPTAD--SDRQGILUOAMLDHQ 220  
 Db 120 ABVRDQFARGDIALAIALQSISLGRCRQ-----QGLYBOAOHFGRS 162  
 QY 146 RREYNE---GILDDDEBEATVTIDVPMWDKVPGALCVTQEGGETEIVRVEAHLKILRK 200  
 QY |:|:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 63 RQYDRARFFQGGLALEPSELEED-WQRI-GALLILEGHYDRYSQASABLLPDYDS 119  
 Db 249 -SSLA-----PDRAQIDETLEBEITPVYLSLGLPGDYAKRKLGSVRNL--W 299  
 QY |:|:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 163 QSLADHORFEPSPSLSRTHQEDQPLRPTILRLAQLPTAD--SDRQGILUOAMLDHQ 220  
 Db 300 SVGGGGASALVGLTREKFMNEAFLR---MTAABQVLDF---VATPSNIPAESFBEVY 351  
 QY |:|:|||:|||:|||:|||:|||:|||:  
 Db 375 EVRAFLD-----DPQAPAPAPEPRPARVIALSGASS-----GTEPGRL-- 413  
 Db 221 GIEGPQGDG--SGLTLDNF--MFLOQIRGVITLQQLIIESEARPS--PAASF--F 271  
 QY 552 TVFSVUDPGVGNVGRDGEPPGVFILEAVRUSSENFTNDYAIRAGVSSESSVDETTVEMVADM 611  
 Db 414 -----LTPVG-----HEARRR-----ANL 428  
 QY 612 LKEASVKLALAAGVAGLTS--LFSQVKPLKSSSFORKDMVSMESDVATIGSVRADS 668  
 Db 429 ISPOAWSAALVALVAFPLGSGWMLTOROSRPEPAPARP--VAPPVOPAPAPIAAPPO 486  
 QY 669 EALPR-----MDARTAENIVSKWOKIKSLSAFLGDPHRIMPLPEUDGRMLKWIWDR 718  
 Db 487 AAMPATAAGDNGNPPDQDQAA-MLKQWQTAKQALGPRERQOMQMLTGSRBVWQK 545  
 QY 719 AARTAOGIVYDTLKLQSVDSITVOSAD--GTRALVENTLESSAAC--SDLPHENPATD 774  
 Db 546 VEQSRQAGEYWKESKLKDIEQY--ADRPRDRAVIAQTVTEVANLYTDNQLRPSRSYD- 601  
 QY 775 VTYTTRTEFVNKSKGWIKTESSVL 799  
 Db 602 -RPYRVRLSVKAPAGNWRIEMKV 625  
 RESULT 9  
 O93AKO PRELIMINARY; PRT; 631 AA.  
 ID O93AKO  
 AC O93AKO;  
 DT 01-DEC-2001 (Tremblel. 19, Created)  
 DT 01-DEC-2001 (Tremblel. 19, Last sequence update)  
 DT 01-MAR-2004 (Tremblel. 26, Last annotation update)  
 DE Cell division protein Ptn2.  
 OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.

Db	582 AVVQAKVDE---VEQVTRGDQLIETRDRDGLVIRYOLVRENNIKLASISLV 630	Db	532 LQSPVVSXK---NNDE-----TSASMPVQLKRN---LGVIGKINDEW---LS 571
RESULT 10		Qy	608 VADMKEASVKILAGAAGVAGLGLISLFSQX----- 636
ID Q8VY16	PRELIMINARY; PRT; 819 AA.	Db	572 QSSLIGRVSVVAL-----LGCTVFFSLKLGSIRSGRLQSMPLSVSARPHSESDFLWKE 626
AC Q8VY16;		Qy	637 ---PLKSSSSFORQDMISS-----MESDVATGSRADSEALPR 673
DT 01-MAR-2002	(TREMBlrel. 20, Created)	Db	627 SGNFRKNUISVNHRNGIVGNIKVILIDMLKMRHCGHDPDALYTKSSQSATLSHSASELHKR 686
DT 01-OCT-2003	(TREMBlrel. 25, Last annotation update)	Qy	674 -MDARTAENIIVSKWQKIKSLAFQPDHRIEMLPFVLGMRKLUWTDRATAETAOGLGLVY 731
DE Hypothetical protein At3g19180.		Db	697 PMTEBAAELVROWENYVGAELGPTHQWSYLSIEVLDLSMVLQW-QTIAQTABAKSCYWR 745
OS Arabidopsis thaliana (Mouse-ear cress).		Qy	732 TLLKLSVNSVTVSAD--GTRALVEATLESRESACLSLPHPENNATDVTYTYTRYEVMSK 788
OC Buxarytae; Viridiplanae; Streptophyta; Embryophyta; Tracheophyta;		Db	746 VLLHLLEVQLAHQI-FEDGJAGEARASIEALLEEAEALIVDESQQPK-NAKYVISTKYIKYILKKQE 804
OC Magnoliophyta; eudicots; rosids;		Qy	789 SG-WKITEFGSV 798
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.		Db	805 DGLWKFCOSSDI 815
OX NCBI_TAXID=3702;		Qy	
[1]	SEQUENCE FROM N.A.	Qy	
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,		Qy	
RA Goldstein A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,		Qy	
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,		Qy	
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,		Qy	
RA Miyanaga M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,		Qy	
RA Seki M., Shinm P., Southwick A., Shinozaki K., Davis R.W., Becker J.R.,		Qy	
RA Theologis A.; Submissions (JAN-2002) to the EMBL/GenBank/DDBJ databases.		Qy	
RL Best Local Similarity 10.6%; Score 431; DB 2; Length 819;		Qy	
DR EMBL; AR071283; ALU66980.1; -.		Qy	
KW Hypothetical protein.		Qy	
SQ SEQUENCE 819 AA; 90750 MW; 6720083F4864657E CRC64;		Qy	
Query Match 10.6%; Score 431; DB 2; Length 819;		Qy	
Best Local Similarity 10.6%; Pred. No. 9; DB 2; Length 819;		Qy	
Matches 200; Conservative 116; Mismatches 311; Indels 224; Gaps 32;		Qy	
Qy 56 SASSSEPAATTATLVLSSPSIRPERIPVPIPDFYQVLAQGQHFLTGIRFAEARFVSK 115		Qy	
Db 81 SRSVSLAATST-----IELPVTCYQOLIGVSVQAEKDEVKSVINLKKT 124		Qy	
Qy 116 PPGFGPSDALISRRQIQQACETLSNPRSRSYNEGLDDEBAAVITDVPDKVPGALC 175		Qy	
Db 125 DREGYETMAAARQDPLMDVRYKLI--LFESYAGVNLKKEPKSPKSPRIPWALGIC 181		Qy	
Qy 176 VEQEGGETBIVLNGEALKERIPLPSFKDQDVWYMAFLDVSRDAMALDPPDFITGEF 235		Qy	
Db 182 LIQEVQGKVLVIDGRALNLNDSPKPYHDIFSMALCAECAINKAFAEVNKVQS-OGPEA 239		Qy	
Qy 236 VEEALKLQIQR-BASSLADPDRQIDCTRLPEIPLPVRULGAPLGDDYAAKRLNGLSGV 294		Qy	
Db 240 LABQOSPLKSKVTRGKLA--LLTQIESLEGGIAPPCTDLGIPRTPEAERRGATAA 297		Qy	
Qy 295 RNLIWMSVGGGASSAVGGLTREK-----PNAEAFLMTMAEQDLP-----VATP 339		Qy	
Db 298 REILQ-----GLSVRASCQIOPWCPCLSDAISRLAATVIVLIPDMLAITRK 346		Qy	
Qy 340 SNTPAESR-----VIEVALVALVAQAFICKKKPHLQDADQFOOLQQAQWAMAEIP 390		Qy	
Db 347 NKQGLESINHQVRVIDPNCPYTMVHLGHIAVGFG-----KONETINKAKTICBLLI 396		Qy	
Qy 391 AMYDTRUNWEIIGLGLCALIGKTYDECRWGL-----DESDSOTRNPAVFEVLEN 446		Qy	
Db 397 A-----SEGVDLKFEEAFCSPFLIKQGSBAEALEKLUQLESNSDAVNRN----IIGK 444		Qy	
Qy 447 SNRDDNDLPLGCKLLETLLAGVTPPRDTK-----DKKKGIGYDPPMVL 494		Qy	
Db 445 ESTS-----ATPSLRAWLMEVNLANTFDTCRCSPLANFFRAREKKPKENKGSSIM 499		Qy	
Qy 495 SY-----LERVEVVOGSPLAAMATMARGAENKASAMQALQKVPSRTDMSAERKD 548		Qy	
Db 500 NHKTNQRQLSTPQFNSS-----QHL-----YTAEQLPTD 531		Qy	
Qy 549 VOTVFSITDPVGAVNGRCGEPGYFIAEVRPSENFTNDYAIHAGVSESSV-DTBTVEMS 607		Qy	
[1]	SEQUENCE FROM N.A.	Qy	
RA STRAIN=PCC6803;		Qy	
RC MEDLINE=97051201; PubMed=8905231;		Qy	
RX Kaneko T., Sato S., Kotani H., Tanaka A., Asanizu E., Nakamura Y.,		Qy	
RA Miyajima N., Hirokawa M., Sugira M., Sasaki S., Kimura T., Okumura S.,		Qy	
RA Hobouchi T., Matsuda A., Murakami N., Naruo K., Yamada M., Yasuda M.,		Qy	
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M.,		Qy	
RA Tabata S.; Sequence analysis of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.		Qy	
RT DNA Ref. 3:09-136(1996); DR EMBL; D63999; RA110060.1; -.		Qy	
RT DR PIR; S76082; S76082.		Qy	
RT RSP; P08622; 18022.		Qy	
DR InterPro; IPR01623; DnaJ_N.		Qy	
DR PRAM; PF00226; DnaJ; 1.		Qy	
DR SMART; SW00271; DnaJ; 1.		Qy	
DR PROSITE; PS0076; DnaJ_2; 1.		Qy	
DR Complete proteome; DR PROSITE; PS0076; DnaJ_2; 1.		Qy	
SQ SEQUENCE 714 AA; 79423 MW; 51B52C16F405ED3B CRC64;		Qy	
Query Match 10.6%; Score 429; DB 2; Length 714;		Qy	
Best Local Similarity 25.1%; Pred. No. 1e-17; Matches 194; Conservative 124; Mismatches 331; Indels 124; Gaps 30;		Qy	
Db 86 IPIDPYQVGAQTHPLTDGIRAFEARVSKPQPGFSSDALT-SRRQIQQACETLSNPRRS 145		Qy	
Db 3 IPIPDYRIGIPIPQSGGETIEQAYDQRLQLPREFSDAATVNLQIAVETLRPEK 62		Qy	
Qy 146 RREVNE--GLIDDE--BATVITDVPDKV-----GALCVLQDGGGERIVLNGEALK 196		Qy	
Db 63 RQADQEWNGMAGDEALGMLRPLTPBLPELCSPB30B3IGLILUDGEBVLYKGVPLHD 122		Qy	
Qy 197 RLPKS--FKQDVVILVMAFLDVSRAAMDPDPDFITQEFEEBA-LJKLQSEGASSLAP 253		Qy	
Db 123 PNPPAGGLEQDYLSVLAHWESELRSRWRQQ-----YEFATASLKLARLQODNDPF 176		Qy	

QY	56	SSSSPATATTATLVLPLPSIDRPERVPIPFDYKQLGAQTHPLTGIRRFAEARVSK	115
Db	177	ALEAEIROLYLRPRTYRTELLAKEGQGE-- QQQGLALLQAMVDRGGIEKGEDYS	233
QY	312	GILTRE-- KPMNAPRLMTAEGDVLFWATPSNPAESFVEVYELVALVAQAFICKKPHL	368
Db	234	GLGNDDFLFIHOLRCHLTVAEONALFL-- PESORPLSVASYLAWHSIMAGVKEDOPMA	291
QY	369	LQDADKQFQQLQAKWAKMAMEIPAMLYDTRNNWEIDFGLERGLCALLIGKVDECMLGLD	428
QY	429	SBDSQYRNPAIEFVLENSNRDDNDLPGLKULETLAGUVFPRPROTKKKLGDDY	488
Db	331	QDP--- PKVA-GLESKATERBDPLIAFTYTFEQMLBEIVPFRDISPETSPKAYF	384
Db	292	IYEAKSLTIQLENCO----- DLAEKVICBILGOTEV- VLAIAID	330
QY	489	DPPMVLSTYLERVWVQ----- GSP-LAAATMARTIGAEHVKASAM-QAQKVPSRYT	539
QY	385	NPSVQVQLQBLPDSFTTDNSFASPALISTATESSETPMVHSSALPDRPLSTVPS--	441
Db	540	DINSAPEKDVQETVFSVPVGNNVRGDRGEPGVTAEAUPSENFTNDYAIRA-----	592
QY	593	-GVSESSVDETTVE-MSVADMLKEASVK----- ILAGVAIGLISLSQK-----	635
Db	492	NGFSSNSAPESTSKHKSPPRKERKRTVKVRFGFLCAGIVGGATLJINRQDPPLG	551
Db	442	-RRGRSPRSRSDVFP----- SADNSSGLAVT-TISPAIAYDTHSLSLGTNGGCDSTS	491
QY	636	----- YFLKSSSSFORKDMVSSMSEDVATGSVRADDSEALPMDARTAENTSKWQK	688
Db	552	LIEDPDVFL----- DQPSPPIPDRATSRNLILSQ-- PNPNQQGQMVQWILD	598
QY	689	IRSLAFGDPHRITMLEMPLEVLDGRMKIKIWTDRAESTAOLGLVYDYLKUSVDSVTS-ADG	747
Db	599	SKKLAEGONYDVGALQSPLAQNLAQQR-AQDQAQKVHQYEHKLQGLAYONPQDP	657
QY	748	TRALVEATLESACSLDVHPENNADYR-TYTRREVWFSKSCWKINGSVL	799
Db	658	NRATVTARVEEISOPPTLGNQOKGKSATKDLYTVQYQLVRHQGVWIKIDQIVV	710
RESULT 12			
Q8LJL2			
ID	Q8LJL2	PRELIMINARY; PRT: 841 AA.	
AC	Q8LJL2;		
DT	01-OCT-2000	(T)EMBL: 15, Created)	
DT	01-OCT-2000	(T)EMBL: 15, Last sequence update)	
DT	01-DEC-2001	(T)EMBL: 19, Last annotation update)	
DE	Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MV11.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; rosids; eudicots II; Brassicales; Brassicaceae; Arabidopsis.		
NCBI_TaxID	3702;		
[1]			
RN			
RP	SEQUENCE FROM N A. MEDLINE=20163099; Nakamura Y.; Asamizu E.; Tabata S.; RX		
RA	"Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequences features of the regions of 4,251,695 bp covered by ninety P1, TAC and BAC clones."; DNA Res. 7:217-221(2000).		
[2]			
SEQUENCE FROM N A. Kameko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.; DR Submitted (SBP:1999) to the EMBL/Genbank/DDBJ databases. EMBL: AP000419; BAB02958.1; -.			
SEQUENCE 841 AA; 93216 MW; 7DBE72618EE8B97 CRC64;			
Best Local Similarity 23.0%; Prod. No. 2.3e-17; Matches 200; Conservative 122; Mismatches 308; Indexes 240; Gaps 35; OC			
QY	56	SSSSPATATTATLVLPLPSIDRPERVPIPFDYKQLGAQTHPLTGIRRFAEARVSK	115
Db	81	SRTSLLAEST----- IELPTVCPQLGVSQAEKOEVSINKTT	124
QY	116	PROFGESDDALISRROJLQACETLSNPSSRSBRYNEGLIDEDEATVTDVPMWDVKGALC	175
QY	125	DABEGYTMEEAARQDLIMDVRKL-- LFESEBYAGNIKEKAPKSPRLPIPWALPLGALC	181
QY	176	VQEGARTEIVRVGAEALKERLKPFSKQDQVUVMALAFLVDYSDAMALDPDPFITYGF	235
Db	182	LIQEVGOKVLUVIDGRALRUDSKPVTIDPLSMALACAAFAEVNKVS-- QGFB	239
QY	236	VERALKULQEE-GASSLAPDRAQIDSTLETPRYVLELGLPLGDYAKRKLNGSLGV	294
Db	240	ILAQSPFLXSYTGLKLA-- ILTQIBESLAPPCTDILGLUPRTPENAAERRGAIAL	297
QY	295	RNTLWSVGGGASALVGGLTREK----- FMNEAFLRMTAABQDVL----- VATP	339
Db	298	RELRLQ----- GLSVEASCQ1QDWPCLQSISRLATEIVDLPWDLAIRT	346
QY	340	SNTIPASFEE----- VVYALVALLYAOFATGKPKHILQDADKOFOOLQAOQKWMAMIP	390
Db	347	NIKSLSHNQRVWIDFCNCFYMWLGHIAVGFSG----- KQNETINKAKTCICL	396
QY	391	AMYDTRANNEDPGLERGLCALLIGKVDEC - RMWGLDSE----- DSQY	434
Db	397	A----- SEQDLKREFAFCSPFLKPLSATAPCTCWIAQGSEAALEKLKQLESNSDSAV	449
QY	435	RNPNAIVEFVLENSNRDDNDLPGLKULETLAGUVFPRPROTK----- DKSP	482
Db	450	RNS----- ILGERSSTS----- ATPSLEAMWLMESULANF PDPTRGCPSLANTFRACKY	499
QY	483	KLGDDYDDPMWLSY----- IERVEVWQGSPLAATMARGAHEVKASAMQALQVFP	536
Db	500	PENKKGSPSPIMKHKTNQRPJSTQFVNSS----- QHL----- LGV	532
QY	537	RTYDRNISARPDKVQETVFSVPVGNNVRGDRGEPGVTAEAUPSENFTNDYATRGYSE	596
Db	533	SVTAVEQLPTDLOSPVYSAK----- NNDE----- TSAMSPSVOLKR----- LGV	574
QY	597	SSV-DETIVNEVSPADMILEKEASVIL----- AAGVAIGLISLF----- SQ	634
Db	575	NKWLDEM----- LSQSSLIGRVSITALUGCTVFFSILKUSGIRSLRQLSMSBISVARSRHSSBD	631
QY	635	KYFLKSSSSFORKDMVSSMSEDVATGSVRA-----	665
Db	632	SEWKTESGNRKNDLSVNRNGI- VENIKVUDMLRKHGCHGEPHDALYLIKSSQSATLS	699
QY	666	--DDSEALPR- MDARTAENIVSKWOKTSKLAGFDRITMLEMPLEVLDGRM----- LK	713
Db	690	68A8SLLKHPMDTEEBLROWENVKAEALPHTVOSLSEVLDMSLVOQESIFLCM	749
QY	714	IWTDRAAETAQIAGLVY-DYTILKUSVTSVAD-- GTRALVEATLESACSLDVIRE	769
Db	750	QW-QTIAQTAAKSCYWRFVLUOAHITFEDGIAGEAABIEALLESEAELVDESQPK	808
QY	770	NNATDVTYTYTREVFWSK3-WKITEGSV	798
Db	809	-NAKYVSTYKRYIYKQDQJWKFCSDI	837
RESULT 13			
Q8DKU7			
ID	Q8DKU7	PRELIMINARY; PRT: 673 AA.	
AC	Q8DKU7;		
DT	01-MAR-2003	(T)EMBL: 23, Created)	
DT	01-MAR-2003	(T)EMBL: 23, Last sequence update)	
DT	01-OCT-2003	(T)EMBL: 25, Last annotation update)	
DR	T1r058 protein.		
GN	OrderedLocusName=tir0758;		
OC	Synechococcus elongatus (Thermosynechococcus elongatus). Bacteria; Cyanobacteria; Chroococcales; Synechococcus.		

OX	NEBI_TAXID=2046;	QY	755 TLERSACSSLVHPENNATDVRYITRVEFWWS
RN	[1]		
RP	SEQUENCE FROM N.A.	PRM;	702 AA.
RC			
STRAIN-BP-1;			
RX	MEDLINE:2222144; PubMed=12240834;		
RA	Watanabe A., Iiguchi M., Kawashima K., Kimura T., Katoh H., Sasaki S.,		
RA	Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,		
RA	Slimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;		
RT	"Complete genome structure of the thermophilic cyanobacterium		
RT	thermosynechococcus elongatus BP-1.";		
RL	DNA Res. 9:123-130 (2001).		
DR	EMBL; AP00571; BAC0809.1; -.		
DR	InterPro; IPR001623; Dnaj_N.		
DR	InterPro; IPR008941; TRR-Like.		
DR	Pfam; PF00256; Dnaj_1.		
DR	SMART; SM00271; Dnaj_1.		
DR	PROSITE; PS0076; Dnaj_2; 1.		
KW	Complete proteome.		
SQ	SEQUENCE 673 AA; 75741 MW; 3427C6B46BBCB83A CRC64;		
Query	Match 10.4%; Score 423; DB 2; Length 673;		
Best	Local Similarity 25.4%; Pred. No. 2.20-17;		
Matches	191; Conservative 105; Mismatches 307; Indels 150; Gaps 28;		
QY	86 IPIDFYQVLGAQTHFLTDGIRRAFEARVSKPPQFGFSDALTSRRQIQAACETLSNRS	145	
Db	3 IPIDFYQVLGVPIQATPQEIQAPRDRLLQLPFTQHSPTIVATRRELIEQAVLREPEQ	62	
QY	146 RREYNE-----GLDDERATVTDVDPKVPICALCVQEGGETEVIRVGEALKL	194	
Db	63 RDAYDHRCTVDPDLIAQDLPRTPHIEISBOLSGALLLYELGNAYQVNVLGDAFL	122	
QY	195 KERLPKSKQ-----DVLVLMALFDYDSRDAHALDPPDFTGCFEFFEALKL	243	
Db	123 KKV-----FERNRKYTPSPAVADLTIVLVALYLGLGIEW-----OQSYESAASDNEAGLVL	177	
QY	244 QEGGASSLAPDIAQDIDETLETPRYVLLGLPLGDDYAAKRLNLGSVURNILWVG	303	
Db	178 QR---VNLPELQBQFQTLNARLARPYRTELLMLPLSD---SANQRGILLURQMSBRGG	232	
QY	304 -GASA-SALVGLTTR--KEPNRKAFLRMTAAEQDTLFVATPSNTPAESTEVYALVAQ	359	
Db	233 IEGRGDDRSGLTVDFLKLQLRSLHLSLTVAQBLF-----BRESRPSAVATLVH	285	
QY	360 AFGKKPHLQD---ADKQFQDQAKUMAMBIAPAMYDINNWEIDFGLERGLCALL	414	
Db	286 ALVARGVHLQPSYCIRKADLQD-----LPHQ-----DVIYLASCILL	326	
QY	415 IGVDECEWMWLGJDSQSYRNPAIVEPVLENSNRDDNDLPGCLKULETWLAGVVFPRF	474	
Db	327 LGQPTE---AALALDHSQDQPTL-DTIRRHAG-BAGDRFLQYVYTQMLTBETIPAF	379	
QY	475 RDTKDKKKFLGDXDPRMVLSYLERVETVOGSISSLAATMARGAERHKASANAOQKV	534	
Db	380 RIDGETPVALEAFADANQTYLAEALSDSTAEPARTA-----419		
QY	535 PSEYDTRNSAEPDQVQETVFSUPVGNNGRDCBPGVIAAVRPSENFETNDYAIRGV	594	
Db	420 -----SALPDEVIRPTV-AVEP-----PLSTACTL-PLDOSRIGQGSLASA	459	
QY	595 SESSVDETTEVMSVADMKEASVK-----ILAGAVA-IGLISLFSQYFLKSS	642	
Db	460 FPPSATATGTSHQSPKRKSFRNCRCAQKQTMFWMGAGVUVLGLG-----AKVYMPAKT	517	
QY	643 SRQDKDMVSSMSBDAVATGSVRADDSEAL---PRMDARTAENIVSKNOKIKLAFGDHR	699	
Db	518 AEAPPPIPPTAPIPVATPTPTPQPTLTLTPEM---ARDQHWTQKQKAQALGRPF	573	
QY	700 IEMLPPEVLDGMKMKIWTDRAEPAOLQVYDYLKULKUSDVSYS-ADGTRALVE---A	754	
Db	574 VDQUTTIAEPSELSRWSRAQGKSESYWVYVILKNEVKETLQLRDREVLYABVEDA	633	
QY			
QY	755 TLERSACSSLVHPENNATDVRYITRVEFWWS	787	
Db	634 RFTVEQGTI-----RDI-SVSDPFRVIFT	656	
RESULT	14		
QY	Q7V0H1 PRELIMINARY;	PRM;	702 AA.
ID	Q7V0H1		
AC	Q7V0H1;		
DT	01-OCT-2003 (TREMBrel. 25, Created)		
DT	01-OCT-2003 (TREMBrel. 25, Last sequence update)		
DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)		
DR	Hypothetical protein.		
DR	Orderedlocusname=PMM1287;		
DR	Prochlorococcus marinus subsp. <i>pastoris</i> (strain CCMP 1378 / MED4).		
OC	Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;		
OC	Prochlorococcus.		
OX	NCBITaxID=59919;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE:2222569; PubMed=12917642; DOI=10.1038/nature01947;		
RA	Rosap G., Larimer F.W., Lauterlin J.E., Malfatti S., Chain P., Ahlgren N.A., Arulano A., Coleman M., Hauser L., Hess W.R., Johnson Z.I., Land M.L., Lindell D., Post A.F., Regalado W., Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A., Webb E.A., Zinser E.R., Chisholm S.W.;		
RA	"Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation.";		
RT	Nature 424:1042-1047(2003).		
RL	EMBL; BX572093; CME197451; -.		
DR	InterPro; IPR001623; Dnaj_N.		
DR	Pfam; PF00256; Dnaj_1.		
DR	SMART; SM00271; Dnaj_1.		
KW	Complete proteome; Hypothetical protein.		
SQ	SEQUENCE 702 AA; 79938 MW; 325A8P3EE74E55BF CRC64;		
Query	Match 8.5%; Score 344.5; DB 2; Length 702;		
Best	Local Similarity 20.5%; Pred. No. 1.5e-12;		
Matches	161; Conservative 15; Mismatches 305; Indels 163; Gaps 31;		
QY	86 IPIDFYQVLGAQTHFLTDGIRRAFEARVSKPPQFGFSDALTSRRQIQAACETLSNRS	145	
Db	3 IPIDFYQVLGVPSATSBEILRAFQLRDKTFDGGFTWVLTQRSELLRLTADLTPES	62	
QY	146 RREYNE-----GLDDERATVTDVDPKVPICALCVQEGGETEVIRVGEALK-----ERLPSK	201	
Db	63 RDAYDHRCTVDPDLIAQDLPRTPHIEISBOLSGALLLYELGNAYQVNVLGDAFL	122	
QY	202 FKDQVVLVMAFLDSRDA---MALDPDPDFTGCFEFFEALKLQERGASSLAPDIAQI	259	
Db	120 READL-----TSDAAQEOQDRSRSYNAQDFOGLQIQLQMGKLG---ERLPSK	172	
QY	260 DENTLEETPRYVLLGLPLGDDYAAKRLNLGSVURNILWVGGSASALVGLTR-----K	317	
Db	173 EEDLVSLLPYRIDLSSRLNDYDHKK---GLSMLENLIKRG-----LEGKNKSYND	225	
QY	318 FMN----EAFIA----MTAASCVLDTWATPENIPANSEFVEVALVAQFIGKPHL	368	
Db	226 FLNQQEFSEFFFOOKIPFLVQDQIDFLFLEQKRGSSSA---GFLAFSLTIAGPARRPK	283	
QY	425 LGLSDSESDQSYRNPAIVEPVLENSNRDDNDLPGCLKULETWLAGVVFPRFRDTKDKFKL	484	
Db	369 LQADAKQFQLOQAKVAME---IPAMLYDTRNNWEIDFGLERGLCALLIGKVDECRM	424	
QY	284 LEARKLKLKLNLSGLOSMPLGCDLID-----VEQSSARFLSSDEKRD	333	
Db	425 LGLSDSESDQSYRNPAIVEPVLENSNRDDNDLPGCLKULETWLAGVVFPRFRDTKDKFKL	484	
QY	334 L-----NNTFGEKLBACIICPKCKNWFLENDLVGYRDLKEDL	371	
QY	485 GOYDPPMVLSYLERVWVQGPMATMARIQABEVKRSAMQALQKVPSRTYNS	544	
Db	372 DSFDFDRIQEFIEQIE-----KKSNTVFKSGPON-----KPIFOQESLKDS	416	
QY	545 EPQDQSTIVPSVD----PVGNNVGRDG----EPCVFIABAVRPSENFETNDYAIRGVSES	598	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 10, 2005, 01:37:44 ; Search time 475 Seconds

Scoring table: BLOSUM62  
GapOp 10.0 , GapExt 0.5

Title: US-10-600-070-2  
Perfect score: 4063

Sequence: 1 MEALSHVGIGLSPFQLCRIP.....YEVFMWSKSGWIKITBGSVLAS 801

Searched: 1710399 seqB, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Published Applications AA,\*

1: /cgn2\_6/ptodata/1/pubpa/US08\_NEW\_PUB.pep:  
2: /cgn2\_6/ptodata/1/pubpa/US07\_PUBCOMB.pep:  
3: /cgn2\_6/ptodata/1/pubpa/US06\_NEW\_PUB.pep:  
4: /cgn2\_6/ptodata/1/pubpa/US06\_PUBCOMB.pep:  
5: /cgn2\_6/ptodata/1/pubpa/US07\_NEW\_PUB.pep:  
6: /cgn2\_6/ptodata/1/pubpa/PCTUS\_PUBCOMB.pep:  
7: /cgn2\_6/ptodata/1/pubpa/US08\_NEW\_PUB.pep:  
8: /cgn2\_6/ptodata/1/pubpa/US08\_PUBCOMB.pep:  
9: /cgn2\_6/ptodata/1/pubpa/US09\_NEW\_PUB.pep:  
10: /cgn2\_6/ptodata/1/pubpa/US09\_PUBCOMB.pep:  
11: /cgn2\_6/ptodata/1/pubpa/US10\_NEW\_PUB.pep:  
12: /cgn2\_6/ptodata/1/pubpa/US10\_PUBCOMB.pep:  
13: /cgn2\_6/ptodata/1/pubpa/US10A\_PUBCOMB.pep:  
14: /cgn2\_6/ptodata/1/pubpa/US10C\_PUBCOMB.pep:  
15: /cgn2\_6/ptodata/1/pubpa/US10D\_PUBCOMB.pep:  
16: /cgn2\_6/ptodata/1/pubpa/US10E\_PUBCOMB.pep:  
17: /cgn2\_6/ptodata/1/pubpa/US10\_FOUR\_PUB.pep:  
18: /cgn2\_6/ptodata/1/pubpa/US10\_NEW\_PUB.pep:  
19: /cgn2\_6/ptodata/1/pubpa/US1A\_PUBCOMB.pep:  
20: /cgn2\_6/ptodata/1/pubpa/US11\_NEW\_PUB.pep:  
21: /cgn2\_6/ptodata/1/pubpa/US60\_NEW\_PUB.pep:  
22: /cgn2\_6/ptodata/1/pubpa/US60\_PUBCOMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No. Score Query Match Length DB ID Description

1 4063 100.0 801 16 US-10-600-070-2 Sequence 2, AppIn

2 4051 99.7 801 16 US-10-600-070-127 Sequence 127, App

3 4051 99.7 801 16 US-10-600-070-129 Sequence 129, App

4 4051 99.7 801 16 US-10-600-070-131 Sequence 131, App

5 4051 99.7 801 16 US-10-739-930-5771 Sequence 5771, App

6 2926 72.0 578 16 US-10-600-070-123 Sequence 123, App

7 1775.5 43.7 760 16 US-10-600-070-125 Sequence 125, App

8 1546 40.5 324 16 US-10-600-070-205 Sequence 206, App

9 1332 32.5 16 US-10-600-070-122 Sequence 122, App

10 1132 27.9 344 15 US-10-424-593-27149 Sequence 27149,

11 973 23.9 364 16 US-10-425-115-266516 Sequence 266516,

## ALIGNMENTS

## RESULT 1

US-10-600-070-2  
; Sequence 2, Application US/10600070

Publication No. US20040139500A1

GENERAL INFORMATION:

APPLICANT: Osteryoung, Katherine W.

APPLICANT: Vlita, Stanislav

APPLICANT: Koksharova, Olga A.

APPLICANT: Gao, Hong

TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of

TITLE OF INVENTION: Use

FILE REFERENCE: MSU-0813

CURRENT APPLICATION NUMBER: US/10/600,070

CURRENT FILING DATE: 2003-06-20

NUMBER OF SEQ ID NOS.: 206

SOFTWARE: PatentIn version 3.2

SEQ ID NO: 2

LENGTH: 801

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-10-600-070-2

Query Match 100.0%; Score 4063; DB 16; Length 801;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEALSHVGIGLSPFQLCRIPATKLRRSHNTTSTICASKNADRLISDFNFTSDSSSS 60  
Db 1 MEALSHVGIGLSPFQLCRIPATKLRRSHNTTSTICASKNADRLISDFNFTSDSSSS 60  
QY 61 PAAATTATLVSPPSPPDRPEREPVIPIDFYQVGAQPHFLGDGIRAFEARVSKPQQG 120  
Db 61 PAAATTATLVSPPSPPDRPEREPVIPIDFYQVGAQPHFLGDGIRAFEARVSKPQQG 120

Qry 121 FSDDALISRRQIQAACTLSNRSRREYNEGLDDERATVITDVKPGALCVIQLQEG 180  
Db 121 FSDDALISRRQIQAACTLSNRSRREYNEGLDDERATVITDVKPGALCVIQLQEG 180  
Qry 181 GETEIVLVRGEALLKERLKPSFKQDVLMALFDVSRRDAMALDPDPRTGFEVREAL 240  
Db 181 GETEIVLVRGEALLKERLKPSFKQDVLMALFDVSRRDAMALDPDPRTGFEVREAL 240  
Qry 241 KUJQBEGASSLAQDPLRAQIDETEITPRYVLEBLIGLPGDDYAKRKLNGLSGRVNLIWS 300  
Db 241 KUJQBEGASSLAQDPLRAQIDETEITPRYVLEBLIGLPGDDYAKRKLNGLSGRVNLIWS 300  
Qry 301 VSGGGASALVGGTREKFMNEFLRMTAEQDLFVATPSNIPAESPEVYEVALLVAQA 360  
Db 301 VSGGGASALVGGTREKFMNEFLRMTAEQDLFVATPSNIPAESPEVYEVALLVAQA 360  
Qry 361 PIKKPHILQDAKQFQOLQAXKUMAMEIPAMYDTRNNEDPGLERGLCALLIGRUE 420  
Db 361 PIKKPHILQDAKQFQOLQAXKUMAMEIPAMYDTRNNEDPGLERGLCALLIGRUE 420  
Qry 421 CRMWLGJLDSQDRYRNPAIVEFVLENSNRDDNDPLGCKULETTLAGVPPFRDTDK 480  
Db 421 CRMWLGJLDSQDRYRNPAIVEFVLENSNRDDNDPLGCKULETTLAGVPPFRDTDK 480  
Qry 481 KFKLGDYDDPMVLSYLVEVVOGSPLAAATMARTGAERYKASAMOLOKVFPSRTD 540  
Db 481 KFKLGDYDDPMVLSYLVEVVOGSPLAAATMARTGAERYKASAMOLOKVFPSRTD 540  
Qry 541 RNSAEPKVOETFSVDPVGNNGRDGPGVTAEARPSENFTENDYAIRGVSESSVD 600  
Db 541 RNSAEPKVOETFSVDPVGNNGRDGPGVTAEARPSENFTENDYAIRGVSESSVD 600  
Qry 601 ETTEVEMSVADMKEASVILAGVAGLISLSQKYLKSSFFORKDMVSMESVATI 660  
Db 601 ETTEVEMSVADMKEASVILAGVAGLISLSQKYLKSSFFORKDMVSMESVATI 660  
Qry 661 GSVRADDSSEALPMDARTAENTVKWOKKLSAFGPDRRIEMPPEVLDGRMKIWTDR 720  
Db 661 GSVRADDSSEALPMDARTAENTVKWOKKLSAFGPDRRIEMPPEVLDGRMKIWTDR 720  
Qry 721 ETRQOLGLVYDYLKLSVDSVTSAKGTRALVETABLESACSLDVPHENNATDVRTYT 780  
Db 721 ETRQOLGLVYDYLKLSVDSVTSAKGTRALVETABLESACSLDVPHENNATDVRTYT 780  
Qry 781 RVEVFWMSKGWKITEGSVLAS 801  
Db 781 RVEVFWMSKGWKITEGSVLAS 801

RESULT 2

US-10-600-070-127  
; Sequence 127, Application US/10600070  
; GENERAL INFORMATION:  
; APPLICANT: Ostryoyoung, Katherine W.  
; APPLICANT: Vitha, Stanislav  
; APPLICANT: Koksharova, Olga A.  
; APPLICANT: Gao, Hongo  
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
; TITLE OF INVENTION: Use  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; CURRENT FILING DATE: 2003-06-20  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: SeqIn version 3.2  
; SEQ ID NO 127  
; LENGTH: 801  
; TYPE: PRT  
; ORGANISM: *Arabidopsis thaliana*  
; US-10-600-070-127

Query Match Score 99.7%; Score 4051; DB 16; Length 801;  
Best Local Similarity 99.8%; Pred. No. 0; Gap 0;

Matches 799; Conservative 0; Mismatches 2; Indels 0; Gap 0;  
Qry 1 MEALSHYQIGLSPFOLCRLLPATTKLRSHNTSTTCASAKWADRLLSDFNITSDDSSS 60  
Db 1 MEALSHYQIGLSPFOLCRLLPATTKLRSHNTSTTCASAKWADRLLSDFNITSDDSSS 60  
Qry 61 FATTATTLVSLPPSITPERHVPIDFYQVLAQTHFLTDGIRRAPEARSKPPQFG 120  
Db 61 FATTATTLVSLPPSITPERHVPIDFYQVLAQTHFLTDGIRRAPEARSKPPQFG 120  
Qry 121 FSDDALISRRQIQAACTLSNRSRREYNEGLDDERATVITDVKPGALCVIQLQEG 180  
Db 121 FSDDALISRRQIQAACTLSNRSRREYNEGLDDERATVITDVKPGALCVIQLQEG 180  
Qry 121 CRMWLGJLDSQDRYRNPAIVEFVLENSNRDDNDPLGCKULETTLAGVPPFRDTDK 480  
Db 121 CRMWLGJLDSQDRYRNPAIVEFVLENSNRDDNDPLGCKULETTLAGVPPFRDTDK 480  
Qry 241 KUJQBEGASSLAQDPLRAQIDETEITPRYVLEBLIGLPGDDYAKRKLNGLSGRVNLIWS 300  
Db 241 KUJQBEGASSLAQDPLRAQIDETEITPRYVLEBLIGLPGDDYAKRKLNGLSGRVNLIWS 300  
Qry 301 VSGGGASALVGGTREKFMNEFLRMTAEQDLFVATPSNIPAESPEVYEVALLVAQA 360  
Db 301 VSGGGASALVGGTREKFMNEFLRMTAEQDLFVATPSNIPAESPEVYEVALLVAQA 360  
Qry 361 PIKKPHILQDAKQFQOLQAXKUMAMEIPAMYDTRNNEDPGLERGLCALLIGRUE 420  
Db 361 PIKKPHILQDAKQFQOLQAXKUMAMEIPAMYDTRNNEDPGLERGLCALLIGRUE 420  
Qry 421 CRMWLGJLDSQDRYRNPAIVEFVLENSNRDDNDPLGCKULETTLAGVPPFRDTDK 480  
Db 421 CRMWLGJLDSQDRYRNPAIVEFVLENSNRDDNDPLGCKULETTLAGVPPFRDTDK 480  
Qry 481 KFKLGDYDDPMVLSYLVEVVOGSPLAAATMARTGAERYKASAMOLOKVFPSRTD 540  
Db 481 KFKLGDYDDPMVLSYLVEVVOGSPLAAATMARTGAERYKASAMOLOKVFPSRTD 540  
Qry 541 RNSAEPKVOETFSVDPVGNNGRDGPGVTAEARPSENFTENDYAIRGVSESSVD 600  
Db 541 RNSAEPKVOETFSVDPVGNNGRDGPGVTAEARPSENFTENDYAIRGVSESSVD 600  
Qry 601 ETTEVEMSVADMKEASVILAGVAGLISLSQKYLKSSFFORKDMVSMESVATI 660  
Db 601 ETTEVEMSVADMKEASVILAGVAGLISLSQKYLKSSFFORKDMVSMESVATI 660  
Qry 661 GSVRADDSSEALPMDARTAENTVKWOKKLSAFGPDRRIEMPPEVLDGRMKIWTDR 720  
Db 661 GSVRADDSSEALPMDARTAENTVKWOKKLSAFGPDRRIEMPPEVLDGRMKIWTDR 720  
Qry 721 ETRQOLGLVYDYLKLSVDSVTSAKGTRALVETABLESACSLDVPHENNATDVRTYT 780  
Db 721 ETRQOLGLVYDYLKLSVDSVTSAKGTRALVETABLESACSLDVPHENNATDVRTYT 780  
Qry 781 RVEVFWMSKGWKITEGSVLAS 801  
Db 781 RVEVFWMSKGWKITEGSVLAS 801

RESULT 3

US-10-600-070-129  
; Sequence 129, Application US/10600070  
; Publication No. US200401395001  
; GENERAL INFORMATION:  
; APPLICANT: Ostryoyoung, Katherine W.  
; APPLICANT: Vitha, Stanislav  
; APPLICANT: Koksharova, Olga A.  
; APPLICANT: Gao, Hongo  
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
; TITLE OF INVENTION: Use  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; CURRENT FILING DATE: 2003-06-20

; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO: 129

; LENGTH: 801

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-10-600-070-129

Query Match 99.7%; Score 4051; DB 16; Length 801;  
Best Local Similarity 99.8%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEALSHVIGLSPFQLCRLPPATTKURRSNHTTICSAKMDRLISDFNTSDSSSS 60

Db 1 MEALSHVIGLSPFQLCRLPPATTKURRSNHTTICSAKMDRLISDFNTSDSSSS 60

QY 61 FATAATTATLVSPPSIDRPERHVPIDPYQVIGAQTGFLTGIRRAFEARVSKPQFG 120

Db 61 FATAATTATLVSPPSIDRPERHVPIDPYQVIGAQTGFLTGIRRAFEARVSKPQFG 120

QY 121 FSDDALISRROQIQAQCTLSNRSRREYNEGLDEEATVITDVKPGALCVQEG 180

Db 121 FSDDALISRROQIQAQCTLSNRSRREYNEGLDEEATVITDVKPGALCVQEG 180

QY 181 GETBIVLRGEALKERIKPSFKQDVVIALMFLDSRDA MALDPDFITGVFVERAL 240

Db 181 GETBIVLRGEALKERIKPSFKQDVVIALMFLDSRDA MALDPDFITGVFVERAL 240

QY 241 KLIQESEGSSLAPDLRAQIDEETLEITPRYVILIGLGLGDDYAKRNLGSVRNLTWS 300

Db 241 KLIQESEGSSLAPDLRAQIDEETLEITPRYVILIGLGLGDDYAKRNLGSVRNLTWS 300

QY 301 VGGGASALVGGTREKMINEARLMTAEQDVLFVATPSNIPAESFVYEVALLVAQA 360

Db 301 VGGGASALVGGTREKMINEARLMTAEQDVLFVATPSNIPAESFVYEVALLVAQA 360

QY 361 FIGKPHILQDAKQFOOLQOAKYMMAEIPAMYDTRNWEIDFLGFLRGCLCALLIGKDE 420

Db 361 FIGKPHILQDAKQFOOLQOAKYMMAEIPAMYDTRNWEIDFLGFLRGCLCALLIGKDE 420

QY 421 CRMWLGJLSESDSOYRNPAIVEFLNSNRDDNDLPGLKLTBTWLAGVVFPRFRDTKD 480

Db 421 CRMWLGJLSESDSOYRNPAIVEFLNSNRDDNDLPGLKLTBTWLAGVVFPRFRDTKD 480

QY 481 KFLKGDYDDPMWLSYLERYVEVVOGSLAAAMARTGAEHVKASAMGALQKVPSRYTD 540

Db 481 KFLKGDYDDPMWLSYLERYVEVVOGSLAAAMARTGAEHVKASAMGALQKVPSRYTD 540

QY 541 RNSAEPKVOETVFSVDVGNNGRDFGEGPVFTABAVERPSENFTNDYAIRAGVSESSWD 600

Db 541 RNSAEPKVOETVFSVDVGNNGRDFGEGPVFTABAVERPSENFTNDYAIRAGVSESSWD 600

QY 601 ETTEVEMSYADMILKEASVKKLAACVAGAIGHLSLSFQKYFLKSSFORKOMVSSMESDVTI 660

Db 601 ETTEVEMSYADMILKEASVKKLAACVAGAIGHLSLSFQKYFLKSSFORKOMVSSMESDVTI 660

QY 661 GSTRADDESEALPMDPARDTAPTAINTSKWOKKLSLAFGPDRIEMPEVLDGRMLKWTDR 720

Db 661 GSTRADDESEALPMDPARDTAPTAINTSKWOKKLSLAFGPDRIEMPEVLDGRMLKWTDR 720

QY 721 ETQOLGLVYDVTUJKLSVSVTSADGTRALVATELERSACSLDVHPENNADVRTYT 780

Db 721 ETQOLGLVYDVTUJKLSVSVTSADGTRALVATELERSACSLDVHPENNADVRTYT 780

QY 781 RYVPMWSKSGWKTEGSQLAS 801

Db 781 RYVPMWSKSGWKTEGSQLAS 801

; GENERAL INFORMATION:  
; APPLICANT: Ostervoung, Katherine W.

; APPLICANT: Vitha, Stanislav

; APPLICANT: Koksharova, Olga A.

; APPLICANT: Gao, Hongo

; TITLE OF INVENTION: Use

; FILE REFERENCE: MSU-08153

; CURRENT APPLICATION NUMBER: US/10/600,070

; NUMBER OF SEQ ID NOS: 206

; SEQ ID NO: 131

; LENGTH: 801

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-10-600-070-131

Query Match 99.7%; Score 4051; DB 16; Length 801;  
Best Local Similarity 99.8%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEALSHVIGLSPFQLCRLPPATTKURRSNHTTICSAKMDRLISDFNTSDSSSS 60

Db 1 MEALSHVIGLSPFQLCRLPPATTKURRSNHTTICSAKMDRLISDFNTSDSSSS 60

QY 61 FATAATTATLVSPPSIDRPERHVPIDPYQVIGAQTGFLTGIRRAFEARVSKPQFG 120

Db 61 FATAATTATLVSPPSIDRPERHVPIDPYQVIGAQTGFLTGIRRAFEARVSKPQFG 120

QY 121 FSDDALISRROQIQAQCTLSNRSRREYNEGLDEEATVITDVKPGALCVQEG 180

Db 121 FSDDALISRROQIQAQCTLSNRSRREYNEGLDEEATVITDVKPGALCVQEG 180

QY 181 GETBIVLRGEALKERIKPSFKQDVVIALMFLDSRDA MALDPDFITGVFVERAL 240

Db 181 GETBIVLRGEALKERIKPSFKQDVVIALMFLDSRDA MALDPDFITGVFVERAL 240

QY 241 KLIQESEGSSLAPDLRAQIDEETLEITPRYVILIGLGLGDDYAKRNLGSVRNLTWS 300

Db 241 KLIQESEGSSLAPDLRAQIDEETLEITPRYVILIGLGLGDDYAKRNLGSVRNLTWS 300

QY 301 VGGGASALVGGTREKMINEARLMTAEQDVLFVATPSNIPAESFVYEVALLVAQA 360

Db 301 VGGGASALVGGTREKMINEARLMTAEQDVLFVATPSNIPAESFVYEVALLVAQA 360

QY 361 FIGKPHILQDAKQFOOLQOAKYMMAEIPAMYDTRNWEIDFLGFLRGCLCALLIGKDE 420

Db 361 FIGKPHILQDAKQFOOLQOAKYMMAEIPAMYDTRNWEIDFLGFLRGCLCALLIGKDE 420

QY 421 CRMWLGJLSESDSOYRNPAIVEFLNSNRDDNDLPGLKLTBTWLAGVVFPRFRDTKD 480

Db 421 CRMWLGJLSESDSOYRNPAIVEFLNSNRDDNDLPGLKLTBTWLAGVVFPRFRDTKD 480

QY 481 KFLKGDYDDPMWLSYLERYVEVVOGSLAAAMARTGAEHVKASAMGALQKVPSRYTD 540

Db 481 KFLKGDYDDPMWLSYLERYVEVVOGSLAAAMARTGAEHVKASAMGALQKVPSRYTD 540

QY 541 RNSAEPKVOETVFSVDVGNNGRDFGEGPVFTABAVERPSENFTNDYAIRAGVSESSWD 600

Db 541 RNSAEPKVOETVFSVDVGNNGRDFGEGPVFTABAVERPSENFTNDYAIRAGVSESSWD 600

QY 601 ETTEVEMSYADMILKEASVKKLAACVAGAIGHLSLSFQKYFLKSSFORKOMVSSMESDVTI 660

Db 601 ETTEVEMSYADMILKEASVKKLAACVAGAIGHLSLSFQKYFLKSSFORKOMVSSMESDVTI 660

QY 661 GSTRADDESEALPMDPARDTAPTAINTSKWOKKLSLAFGPDRIEMPEVLDGRMLKWTDR 720

Db 661 GSTRADDESEALPMDPARDTAPTAINTSKWOKKLSLAFGPDRIEMPEVLDGRMLKWTDR 720

QY 721 ETQOLGLVYDVTUJKLSVSVTSADGTRALVATELERSACSLDVHPENNADVRTYT 780

Db 721 ETQOLGLVYDVTUJKLSVSVTSADGTRALVATELERSACSLDVHPENNADVRTYT 780

; RESULT 4  
; Sequence 131, Application US/10600070  
; Publication No. US20040139500A1

QY 781 RVEVFVWSKGWKTTEGSYLAS 801  
Db 781 RVEVFVWSKGWKTTEGSYLAS 801

RESULT 5  
US-10-739-930-5771  
; Sequence 5771, Application US/10739930  
; Publication No. US2004016190A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovacic, David K.  
; TITLE OF INVENTION: NUCLEAR ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 3B-21(53377)B  
; CURRENT APPLICATION NUMBER: US/10/739, 930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 5771  
; LENGTH: 801  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE: OTHER INFORMATION: Clone ID: ARA8H-23APR03-C13643\_1.p  
; US-10-739-930-5771

Query Match 99.7%; Score 4051; DB 16; Length 801;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRLASHVGIGSPFQLCRLLPPTKLRSHTSTTCASKRDRUDSDFTSDSSSS 60  
Db 1 MRLASHVGIGSPFQLCRLLPPTKLRSHTSTTCASKRDRUDSDFTSDSSSS 60

QY 61 FTTATTATLVSPLPSDRPERHPVPIPDYQVLAQTHPLTDGIRKAPEARSKPPQFG 120  
Db 61 FTTATTATLVSPLPSDRPERHPVPIPDYQVLAQTHPLTDGIRKAPEARSKPPQFG 120

QY 121 FSDDALISRQIQLQACETLSNRSRBSRYNEGLIDDEBATVITDVPNDKVGALCVQBG 180  
Db 121 FSDDALISRQIQLQACETLSNRSRBSRYNEGLIDDEBATVITDVPNDKVGALCVQBG 180

QY 181 GETTIVLRVGEALKERLKPSPKFDVVLVMAFLDVSDAMALDPDPDFTYEFVEAL 240  
Db 181 GETTIVLRVGEALKERLKPSPKFDVVLVMAFLDVSDAMALDPDPDFTYEFVEAL 240

QY 241 KLUQEEGASSLARDLRAQIDEETEPRVYLELGLPGLDDYAAKRLNGSGVRNLMS 300  
Db 241 KLUQEEGASSLARDLRAQIDEETEPRVYLELGLPGLDDYAAKRLNGSGVRNLMS 300

QY 301 VGGGGASALVGGITREKFMEFLRMTAEQDLPVATPSNIPAESFVEVVALVAQA 360  
Db 301 VGGGGASALVGGITREKFMEFLRMTAEQDLPVATPSNIPAESFVEVVALVAQA 360

QY 361 FIGKKPHILQDAKQFOOLQOKAMAMIPAMLYDTRNNWEIDPLGLCKLUEFTWLAGVWPRFRDK 420  
Db 361 FIGKKPHILQDAKQFOOLQOKAMAMIPAMLYDTRNNWEIDPLGLCKLUEFTWLAGVWPRFRDK 420

QY 421 CRMWLGIDSESDSQYRNPAIETVTLSENRRDDMDPLGCKLUEFTWLAGVWPRFRDK 480  
Db 421 CRMWLGIDSESDSQYRNPAIETVTLSENRRDDMDPLGCKLUEFTWLAGVWPRFRDK 480

QY 481 KFKLGDYDPPMVLSYLERVEVQGSLAATMARGAEHKVAKSAMQALQKVPSPRYTD 540  
Db 481 KFKLGDYDPPMVLSYLERVEVQGSLAATMARGAEHKVAKSAMQALQKVPSPRYTD 540

QY 541 RNSAEPKVQOETVFSVPVGNNGRDGPQGPVTAEAVRPSENFTNDYAIRAGVSESSVD 600  
Db 541 RNSAEPKVQOETVFSVPVGNNGRDGPQGPVTAEAVRPSENFTNDYAIRAGVSESSVD 600

QY 601 ETTVENVSADMKEASTKILAAAGVAGLISLISQKTYKSSSFORKDMVSMSESVATI 660  
Db 601 ETTVENVSADMKEASTKILAAAGVAGLISLISQKTYKSSSFORKDMVSMSESVATI 660

QY 661 GSVARDDSEALPRMDARTAENIVSKMOKIKSLAFGDPHRIMLPEVLDGRMLKIWTDA 720  
Db 661 GSVARDDSEALPRMDARTAENIVSKMOKIKSLAFGDPHRIMLPEVLDGRMLKIWTDA 720

QY 721 ETQOLGLVYDYTLIKUSVDSVTVSADGTRALVEATELEFASACSLVHPPENNATDVRYT 780  
Db 721 ETQOLGLVYDYTLIKUSVDSVTVSADGTRALVEATELEFASACSLVHPPENNATDVRYT 780

QY 781 RVEVFVWSKGWKTTEGSYLAS 801  
Db 781 RVEVFVWSKGWKTTEGSYLAS 801

RESULT 6  
US-10-600-070-123  
; Sequence 123, Application US/10600070  
; Publication No. US2004013950A1  
; GENERAL INFORMATION:  
; APPLICANT: Ostrovny, Katherine W.  
; APPLICANT: Vitka, Stanislav  
; APPLICANT: Koksharova, Olga A.  
; APPLICANT: Gao, Hongbo  
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; CURRENT FILING DATE: 2003-06-20  
; NUMBER OF SEQ ID NOS: 206  
; SEQ ID NO 123  
; LENGTH: 578  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; US-10-600-070-123

Query Match 72.0%; Score 2926; DB 16; Length 578;  
Best Local Similarity 99.8%; Pred. No. 2.8e-237;  
Matches 577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 79 RPERHPVPIPDYQVLAQTHPLTDGIRKAPEARSKPPQFGFSDALISRQIQLQACE 138  
Db 79 RPERHPVPIPDYQVLAQTHPLTDGIRKAPEARSKPPQFGFSDALISRQIQLQACE 138

QY 139 TISNPSPRERNEGLIDDEAVTITDVPWDKVGALCVLDRGGERBIVLURVGEALKERL 198  
Db 139 TISNPSPRERNEGLIDDEAVTITDVPWDKVGALCVLDRGGERBIVLURVGEALKERL 198

QY 61 TISNPSPRERNEGLIDDEAVTITDVPWDKVGALCVLDRGGERBIVLURVGEALKERL 120  
Db 61 TISNPSPRERNEGLIDDEAVTITDVPWDKVGALCVLDRGGERBIVLURVGEALKERL 120

QY 199 PRSKFDQVVLMALFDVSDAMALDPDPDTGTYEFVEALKLUQEEGASSLAPDRAQ 258  
Db 199 PRSKFDQVVLMALFDVSDAMALDPDPDTGTYEFVEALKLUQEEGASSLAPDRAQ 258

QY 121 PRSKFDQVVLMALFDVSDAMALDPDPDTGTYEFVEALKLUQEEGASSLAPDRAQ 180  
Db 121 PRSKFDQVVLMALFDVSDAMALDPDPDTGTYEFVEALKLUQEEGASSLAPDRAQ 180

QY 259 IDETFBETPRVYLELGLPGLDDYAAKRLNGSGVRNLMSVGGGGASALVGGITREKF 318  
Db 259 IDETFBETPRVYLELGLPGLDDYAAKRLNGSGVRNLMSVGGGGASALVGGITREKF 318

QY 181 IDETFBETPRVYLELGLPGLDDYAAKRLNGSGVRNLMSVGGGGASALVGGITREKF 240  
Db 181 IDETFBETPRVYLELGLPGLDDYAAKRLNGSGVRNLMSVGGGGASALVGGITREKF 240

QY 319 MNEAFRMTAAQDIFVATSNIPAESFVEVVALVAQA 360  
Db 319 MNEAFRMTAAQDIFVATSNIPAESFVEVVALVAQA 360

QY 241 MNEAFRMTAAQDIFVATSNIPAESFVEVVALVAQAFIGKKPHILQDAKQFOO 300  
Db 241 MNEAFRMTAAQDIFVATSNIPAESFVEVVALVAQAFIGKKPHILQDAKQFOO 300

QY 379 LQOKAYNAMEIPAMLYDTRNNWEIDPLGLCKLUEFTWLAGVWPRFRDK 438  
Db 379 LQOKAYNAMEIPAMLYDTRNNWEIDPLGLCKLUEFTWLAGVWPRFRDK 438

QY 301 LQOKAYNAMEIPAMLYDTRNNWEIDPLGLCKLUEFTWLAGVWPRFRDK 360  
Db 301 LQOKAYNAMEIPAMLYDTRNNWEIDPLGLCKLUEFTWLAGVWPRFRDK 360

QY 439 IVEFVLENISNEDNDLPLGCKLUEFTWLAGVWPRFRDTDKKKFLGDDYDPMVLSYE 498  
Db 439 IVEFVLENISNEDNDLPLGCKLUEFTWLAGVWPRFRDTDKKKFLGDDYDPMVLSYE 498

QY 361 IVEFVLENISNEDNDLPLGCKLUEFTWLAGVWPRFRDK 420  
Db 361 IVEFVLENISNEDNDLPLGCKLUEFTWLAGVWPRFRDK 420

QY 499 RVEVQGSPLAATMARGAEHKVAKSAMQALQKVPSPRTRNSAEPKVQOETVFSVD 558  
Db 499 RVEVQGSPLAATMARGAEHKVAKSAMQALQKVPSPRTRNSAEPKVQOETVFSVD 558

QY 421 RVEVQGSPLAATMARGAEHKVAKSAMQALQKVPSPRTRNSAEPKVQOETVFSVD 480  
Db 421 RVEVQGSPLAATMARGAEHKVAKSAMQALQKVPSPRTRNSAEPKVQOETVFSVD 480

QY 559 VGNVGRDGEPGVFIABAVPSENFTNDYAIRAGVSESSVDETTEVMSADMKEASVK 618

RESULT 7  
US-10-600-070-125

; Sequence 125, Application US/10600070  
; Publication No. US20040139500A1  
; GENERAL INFORMATION:  
; APPLICANT: Ostryoyoung, Katherine W.  
; APPLICANT: Gao, Hongo  
; APPLICANT: Vitha, Stanislav  
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
; TITLE OF INVENTION: Use  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; CURRENT FILING DATE: 2003-06-20  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO: 125  
; LENGTH: 760  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; US-10-600-070-125

Query Match 43.7%; Score 1775.5; DB 16; Length 760;  
Best Local Similarity 47.9%; Pred. No. 4.1e-140;  
Matches 390; Conservative 119; Mismatches 213; Indels 93; Gaps 16;

Qy 12 SPFOLCLRUPATKLRRSHNTSTTC-SASKWADRLLSDPNF---TDSSSSSPATAT 65  
Db 14 APPASPLSPRPRPRRRPPPHPSACRAANRWAERLFDHFLPRAESDPPSPAPAPA 73

Qy 66 TTATVLULPPSISDRPERHPVPIIDFYQVLAQTHFLTGIRRAFEARVSKKPQFGFSDDA 125  
Db 74 APSASPFPVPLFDPDAERSLPLQDVFYKVLGAEPHFLGDFGTRAFERIAKPOGYYSTDA 133

Qy 126 LISRRQLIQACETLSNPSRSRREYNEGLDDEBATTITDVPMDKXPAVGALCVLQREGGETE 185  
Db 134 LVGRQRQLQIANDTLNQNSRTOYDRASENREALTDIANDR-----EAGEALA 184

Qy 186 VLRVGBALLKLERPKSKFKDQVULVMAFLDVSRDAMALDPPDTGTYRFVERALKLQD 245  
Db 185 VLVGTGEQQLDREPDKPRFKDQVULVMAFLDVSRDAMALSPPDYTGCCVLERALKLQD 244

Qy 246 EGASSAIPDLRQAIDETLEBTPRYVIELGLPGLDDYAAKRNLGSURNILMSVGCG 305  
Db 245 DCASNAPDLSQDPTERUBITRCVSLSLPDTIEHKKGQBGQGLGARNILMSVGCG 304

Qy 306 ASALVGGLTREKTMNEAFLRMTAEQDVLFVATPSNTIPASFSEVEYEVALLVAQAFIGK 365  
Db 305 IATVGGGSFSEAPMNEAFLRMTSIBOMDPFSKTPNSIPPEWEIYINVALAHVAQAIISK 364

Qy 366 PHILQDADKQKQLOQAKWMMAMEIPAMLYDRTRNNWELPGLERGICALLIGKVDCRMML 425  
Db 365 PQPIMMADDLFRQOKENVFS---HYADN---EMDIALERAFCSLIVGVDYKSCRMWL 416

Qy 426 GLDSEDSQYRNPAIVEFLBNSN-RDDNDLPLGJCKLUEITLACWVFFPRFRDIKOKKEKL 484  
Db 417 GIDNESSPYRDPKTLETFITVNTSISSEENDLPGJCKLUEITLACWVFFPRFRDIKOKKEKL 476

Qy 485 GDYYDDPMWLSYLERVEVQGSPPLAAMTARGA-----HYKASAMQALQKVF-SR 537  
Db 477 GDYDIDPFLVSLYLERMEGGASHALAAAIAKGGAQATAALGTVKSNALQAFNKFVPLIE 536

Qy 538 YTDNRNSAEPKDVQYETVTISVDPVGNNGRDPYEPVTAEVRPSPNPEFTNDYAIRGVSS 597  
Db 537 QDRS3D-----KDG-PGCVL-----ENPDO----- 559

RESULT 8  
US-10-600-070-206

; Sequence 206, Application US/10600070  
; Publication No. US20040139500A1  
; GENERAL INFORMATION:  
; APPLICANT: Ostryoyoung, Katherine W.  
; APPLICANT: Vitha, Stanislav  
; APPLICANT: Gao, Hongo  
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
; TITLE OF INVENTION: Use  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; CURRENT FILING DATE: 2003-06-20  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO: 206  
; LENGTH: 324  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; US-10-600-070-206

Query Match 40.5%; Score 1646; DB 16; Length 324;  
Best Local Similarity 100.0%; Pred. No. 8.6e-130;  
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEALSHVGIGLSPQOLCRUPATKLRRSHNTSTTC-SASKWADRLLSDPNFNTSDSSSS 60  
Db 1 MEALSHVGIGLSPQOLCRUPATKLRRSHNTSTTC-SASKWADRLLSDPNFNTSDSSSS 60

Qy 61 PATATTATLULPPSISDRPERHPVPIIDFYQVLAQTHFLTGIRRAFEARVSKKPQFG 120  
Db 61 PATATTATLULPPSISDRPERHPVPIIDFYQVLAQTHFLTGIRRAFEARVSKKPQFG 120

Qy 121 PSDLALSRQDQLQACETLSNPSRSRREYNEGLDDEBATTITDVPMDKXPAVGALCVLQ 180  
Db 121 PSDLALSRQDQLQACETLSNPSRSRREYNEGLDDEBATTITDVPMDKXPAVGALCVLQ 180

Qy 181 GETBIVLURVGBALLKLERPKSKFKDQVULVMAFLDVSRDAMALDPPDTGTYRFVERALKLQD 240  
Db 181 GETBIVLURVGBALLKLERPKSKFKDQVULVMAFLDVSRDAMALDPPDTGTYRFVERALKLQD 240

Qy 241 KLUQERGASSAIPDLRQAIDETLEBTPRYVIELGLPGLDDYAAKRNLGSURNILMSVGCG 300  
Db 241 KLUQERGASSAIPDLRQAIDETLEBTPRYVIELGLPGLDDYAAKRNLGSURNILMSVGCG 300

Qy 301 VEGGGASALVGLTREKFMNNEFL 324  
Db 301 VEGGGASALVGLTREKFMNNEFL 324

RESULT 9  
US-10-600-070-122

; Sequence 122, Application US/10600070  
; Publication No. US20040139500A1

: GENERAL INFORMATION:  
; APPLICANT: O'reillyoung, Katherine W.  
; APPLICANT: Vitha, Stanislav  
; APPLICANT: Koksharova, Olga A.  
; APPLICANT: Gao, Hongo  
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; CURRENT FILING DATE: 2003-06-20  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 122  
; LENGTH: 525  
; TYPE: PRT  
; ORGANISM: *Oryza sativa*  
; US-10-600-070-122

Query Match 32.8%; Score 1332; DB 16; Length 525;  
Best Local Similarity 48.8%; Pred. No. 5, 4e-103; Matches 287; Conservative 88; Mismatches 137; Indels 76; Gaps 11;

Qy 81 ERHVPIDPYQVIGAQTHPLTDGIRAFEARNSKPKPQFGPSDALSRSRQTLQACETL 140  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||:  
3 ERSPLQDQFDYKVTLGASPHFLSGCIRRAFEARIKAPPQGYSTDALVRGRQMLQIAHDTL 62

Qy 141 SNFRSRBRYNEGLIDDERATVITDVPMDKPGALCVLQEGGSTEVIRVGALLKERLPK 200  
|||:|||:|||:|||:|||:|||:|||:  
63 MNQNSRRTQYDRAISENREBALTMIDAKD-----EAGELAVLVTGECOLLLRPPK 113

Qy 201 SPKQDVVILVMAFLDSRDAMALDPDPFITYEFVRAALKLQOESEGASSLAPDLRQID 260  
|||:|||:|||:|||:|||:|||:|||:  
114 RPKQDVVILVMAFLDSRDAMALSPDPDVIGCCEVLERALKLQEDGASNLPDLQSID 173

Qy 261 ETLEETTPRYVLEELGLPLGDDYAAKINGLISGVURNILWSVGGGASALVSGLTREKFMN 320  
|||:|||:|||:|||:|||:|||:  
174 ERLBEITTPRCVHLILSLPDPDTEHHKKRQBGARNLMSVGRGGIATVGGGSRFAFMN 233

Qy 321 EAFLRMTAAEQLDFLWATPSNIAESESEVYEAVALVAQFIGKKPHILQDADKQPOQLQ 380  
|||:|||:|||:|||:|||:|||:  
234 EAFLRMTSIEQMDPFSKTPNSIPPEWEIYINVALHQAQITSKRQPFIMMADPLRQO 293

Qy 381 OAKVAMAEIPAMYLHYDRNWEIDLPGLEGLCALLIGYDECMWGLSDSESOYRNIAIV 440  
|||:|||:|||:|||:|||:  
294 KENIGS---HYADN---EMDALERAFCSILVNGDVSCKRMWLGIDNESSPYRDKIL 345

Qy 441 EYVLLESN-RDNDPLDPLGCKLKTETWLAGWVERPRFRKDKEKFKLQGDYDPWMWSTLER 499  
|||:|||:|||:|||:|||:|||:  
346 EPTVNTSSISEENDLPLPGCKLKTETWLFPEVFRSRDRGMRQPLRGLDYYDDEPVEVSLYER 405

Qy 500 VEVVOGSPLAAMTMRIGAE---HVKA-SAMQLOQKVF- SRYTDNRNSAEPKVOET 552  
|||:|||:|||:|||:  
406 MEGGGASHLAAIAAIKGAQATAALGTVKSNIAOFNKVFFLIEQDRSAMANT---- 460

Qy 553 VPSVDPVGNNGRDBGEPCVFLAAVPRPSNFENDYATRAGVSBSSVDETVMVADM 612  
|||:|||:  
461 -----KDG-PGGYL-----ENFO-----ENPAHDS 481

Qy 613 XEASVKILAAVGVAIGLISLPSOKYFLKSSSFQRDKMVSSMEDVATI 660  
|||:|||:  
482 RNAALKITISAGALPALLAVIGAKY-----LPRKRLSAIRSBRHSV 522

RESULT 10  
US-10-424-599-271849  
; Sequence 271849, Application US/10424599  
; Publication No. US20040031107A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovacic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 266516  
; LENGTH: 364  
; TYPE: PRT  
; FEATURE:  
; ORGANISM: *Zea mays*  
; OTHER INFORMATION: Clone ID: MRT4577\_174665C.1.pep  
; US-10-425-115-266516

Query Match 23.9%; Score 973; DB 16; Length 364;  
Best Local Similarity 54.8%; Pred. No. 5, 4e-73; Matches 198; Conservative 54; Mismatches 79; Indels 30; Gaps 4;

Qy 12 SPFQLCRLPATPKLRRSHNT---SITICSAKWDRLISDFNF-----TSDSS 57  
|||:|||:  
Db 14 APPASFLP---LRURSHRPPPGSTCRAASRWRDPLRFLPHLPAAPPMASSS 68

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-15(5323)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 271849  
LENGTH: 344  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_87499C.1.pep  
US-10-424-599-271849

Query Match 27.9%; Score 1134; DB 15; Length 344;  
Best Local Similarity 66.4%; Pred. No. 1, 3e-86; Matches 223; Conservative 48; Mismatches 53; Indels 12; Gaps 3;

Qy 16 LQRLLPATT-----KLRS--HNNTTTCASWKADRUDSFNITSDDSSSSFATT 66  
|||:|||:  
Db 12 LCPHPHTTHPTPKPNLRSLSRGASLSATSKWMLIADFQGDAAS---TSTS 68  
; ISRQLQACETLSNRSRBYNEGLIDDERATVITDVPMDKPGALCVLQEGGSTEVIRVGALLKERLPK 200  
; ISRQLQACETLADPTTSRREYNQSLVDDEBAILTOPIFDKPGALCVLQEGGSTEVIRVGALLKERLPK 188  
; LAVGEALLKERLPKSFQDVVILVMAFLDVSRDAMALDPDPFITYEFVRAALKLQOE 246  
; ISRQLQACETLSNRSRBYNEGLIDDERATVITDVPMDKPGALCVLQEGGSTEVIRVGALLKERLPK 186  
; ISRQLQACETLADPTTSRREYNQSLVDDEBAILTOPIFDKPGALCVLQEGGSTEVIRVGALLKERLPK 188  
; LRGQGLRERLPIKTFQDVVILVMAFLDVSRDAMALSPDFIACEMLRALKLQOE 248  
; GASSLADDLRAQIDETLEETPRYVLEELGLPLGDDYAAKINGLISGVURNILWSVGGGA 306  
; GATSLADDLRAQIDETLEETPRCIVBLLLALPDDSHARRBREGGLGVURNILWSVGGGA 308  
; SALVGALTREKEMNEAFLRMTAAEQLDFLWATPSNI 342  
; ATARGETTREDMEAFPLHTMAEAOVLFVATPSTI 344

QY 58 SSSPATTATTLVSLPPSIDRPRHVPIDPVQVIGAQTHITLDGIRRAFFARVKP 117  
Db 69 SSSP-----VPIPEADRALPLPDKTILGSRPHLGDRGIRAFRIAPP 117  
QY 118 QFQFSDDALISRDQLQOACETTSNPRSRREYNGLDDBEATVTDWPKDQFGALCVL 177  
Db 178 QSGGETETIVLVRORALLERLPLPSFKDQVVLVVALAFDVSRAWALUPPDPTGYEPRV 237  
QY 178 QKGEAOLVLTAEGBHLQDPRPKFQDQVVLAMALAYDISRLAWAASPPDVICCEVLR 237  
Db 238 EALKUJQBERGASSLAQDRAQIDTELETPRIVLLELGLPLADDYAAKRNLNSGVNT 297  
Db 238 RAKLUQDQGASNLAPELPSQINTELETPRIVLLELALPTBKHKQRSQGAGNI 297  
QY 298 LWSVGGGASALVGLTRBKFMNAFLRMAAFLRMAAFLQDFLWATPNTIPASPEVVALV 357  
Db 298 LWSVGGRGGIATVCGPSRRAFMNBAFLQMTSALQMDFFSKTPNISPPWEIISVALVH 357  
Qy 358 A 358  
Db 358 A 358

RESULT 12  
US-10-437-963-172416  
; Sequence 172416, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/437, 963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO: 172416  
; LENGTH: 480  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1): (480)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_70554C.1.pep  
US-10-437-963-172416

Query Match 23.9%; Score 972; DB 16; Length 480;  
Best Local Similarity 42.0%; Pred. No. 1e-72; Mismatches 147; Indels 88; Gaps 15;  
Matches 227; Conservative 79; Mismatches 147; Indels 88; Gaps 15;

QY 286 KKLNLGSVNTLWSVCGGASALVGGTREKTMNEFLRMTAAEQDULFWATPSNPAE 345  
Db 3 KRGQGLOCARNTLWSVGRGGIATVCGPSRRAFMNBAFLRMTSIEQMDPFSKTPNISPP 62  
QY 346 STEVVEVALALVQAFIGKPHILQDAKQFOQQLQAKMAMEIPAMLYDTRNWEDPG 405  
Db 63 WPEIYVVALAHVQAIIKSQRPQIMMADLFLQKFNIGS---HYAQN---ENDLA 114  
QY 406 LERGLCALLIGKTYDECRLMGLUSESDQYRNPAIVEFVLENS-RDDNDLPGICKLLET 464  
Db 115 LERAPCSLIVGDVSKCRMWLGIDNESSPYREPANLFIVTNSSISEENDLFLPGICKLLET 174  
QY 465 WLAGWVPPFRDITKDKKEPKLGDYDTPDPMVLSYLERVEVUVQGSPLAATMARIAGAE--- 521

RESULT 13  
US-10-425-115-242115  
; Sequence 242115, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425, 115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO: 242115  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_152391C.1.pep  
US-10-425-115-242115

Query Match 21.1%; Score 856; DB 16; Length 416;  
Best Local Similarity 40.7%; Pred. No: 4.9e-63; Mismatches 82; Mismatches 126; Indels 70; Gaps 11;  
Matches 191; Conservative 82; Mismatches 126; Indels 70; Gaps 11;

QY 348 EVVVALALVQAFIGKPHILQDAKQFOQQLQAKMAMEIPAMLYDTRNWEDPG 407  
Db 3 EIVSVAPPRAQIAKSPRQIMMADLFLQKFNIGS---YPYEN---EMDIALE 54  
QY 408 RGICALLIGKTYDECRLMGLUSESDQYRNPAIVEFVLENS-RDDNDLPGICKLLET 466  
Db 55 RAUCSLSLVGDICNCRMWLGIDNESSPYREPANLFIVTNSSISEENDLFLPGICKLLET 114  
QY 467 AGVYPPFRDITKDKKEPKLGDYDTPDPMVLSYLERVEVUVQGSPLAATMARIAGAE--- 520  
Db 115 VSEVVPFSRDTGTOFLGDKYDTPDPMVLSYLERVEVUVQGSPLAATMARIAGAE--- 174  
QY 521 HVQGSMQALQKUFPSPRYTDRNAEPAKDVQETVFSVPVGNNGRGEPPGVIAEVPRS 580  
Db 175 TVSSAQMFSKVFP-----LIEQDMSGKQI-----PS 203  
QY 581 ENFTNDYAIRGVSESSVDET'VEMSVAIDLKEASYKILAGVAIGLISLSQKVFLKS 640

Db 204 DDLDKS---LEKLAQGSVAGDILHDSRNDL---KLISAGTLLAFAVIGLKICPRN 254  
 Qy 641 SSSPQRKDMSMNSEDVATIGSVTRADSEA-----LPRMDARTAENIVSKOKIKSLA 693  
 Db 255 KS-----LPALRGEBGVSAVVDSDGPADEEPELIPRMDRKLADIVRKWQSKA 307  
 Qy 694 FGPDPHRLEMPLPVLDGEMLKIKWTDRAETAQIQLGVVYTLKLSVSDVTSVSGTRALVE 753  
 Db 308 LGFBEHTVTALQEVTGNGNMKWTDRABERHGFWEYALSGVTSITVSNGRRAVE 367  
 Qy 754 ATIBESACSLSDLVHPENNATDVRTYTYTRIEVETSK-SGWKITESVULAS 801  
 Db 368 ATIEEVGRITDVTDPKNDAYDPTKTYARVEMTSRPAGWRITERGAVLKs 416  
 RESULT 14  
 US-10-600-070-167  
 ; Sequence 167, Application US/10600070  
 ; Publication No. US20040139500A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Osteryoung, Katherine W.  
 ; APPLICANT: Vitha, Stanislav  
 ; APPLICANT: Koksharova, Olga A.  
 ; APPLICANT: Gao, Hongo  
 ; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of  
 ; CURRENT APPLICATION NUMBER: US/10/600, 070  
 ; CURRENT FILING DATE: 2003-06-20  
 ; NUMBER OF SEQ ID NOS: 206  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO: 167  
 ; LENGTH: 768  
 ; TYPE: PRT  
 ; ORGANISM: Nostoc punctiforme  
 ; Query Match Best Local Similarity 24.9%; Pred. No. 1.1e-34; Length 768;  
 ; Matches 212; Conservative 128; Mismatches 286; Indels 227; Gaps 28;  
 Qy 84 VPIPIDPVQVLGAQTHFLDGIRRAFEARVSPKQFGSSDDALLSRISRQLQACETISNP 143  
 Db 1 VPIPIDPVQVLGAQTHFLDGIRRAFEARVSPKQFGSSDDALLSRISRQLQACETISNP 60  
 Qy 144 RSRREYNEGLL-----DDEARAVIT-----DYPWDKVPGALCVLQ 178  
 Db 61 KORTSYDOLYLAKAYDPONLAMALAVAGENRTTESTKRSGSDTOSLIGETQDENGVALILQ 120  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(5323)B  
 ; CURRENT APPLICATION NUMBER: US/10/424, 599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 28564  
 ; SEQ ID NO: 177901  
 ; LENGTH: 204  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRP3847\_131660C.1.pep  
 ; US-10-424-599-177901  
 ; Query Match Best Local Similarity 12.3%; Score 499; DB 15; Length 204;  
 ; Matches 104; Conservative 39; Mismatches 58; Indels 2; Gaps 2;  
 Qy 221 -----ANALDPPDTITGVFVEEALKLQQEGASSLAPDLRAQIDETIEBTPRYV 271  
 Db 181 WQGHYNAASIDE-----TGQB-----LIVREG--LFSSTIQABIQDLYKRRYI 225  
 Qy 272 LELLGFLGDDYNAKRNLGSGVTLWSVGG-GGASALVGSGITRE--KEMNEARFLMT 327  
 Db 226 LELLALP--QEKTAERSQGLELQNLBDRGIGDTNTDESGNANIIDPLRFQCOLRHLI 283  
 Qy 328 AAEQVQLFVATPSNIPAESFEEVYEAVALVAGQAFIGKPKHLCDAKQFQOLQOQKUMAM 387  
 Db 284 VAFOHKLPEAQSKR--SSAVATYLVALLIARGAQPQALPQMLVRLGKQ----- 337  
 Qy 388 EIPAMLYDTRNNWBDIGLERGICALIGKVDCECMWLGIDSQNSRNPATEVPLENS 447  
 Db 338 -----DYLHQSPCALIGQTEATRVTLSOB--YE--ALAFIREKS 376  
 Qy 448 NRDDNDLUPGLCKULETWLAGVVFPRFDTKDKFKKGDDYDDPMVTSYLERVEVVOGSP 507  
 Db 377 -QDSPDLPGLCIXYAOQWLOHEUFPFRDLQANQAFKDYFANQVOYLE----- 426  
 Qy 508 LAATATMARIIGHEVKASAMQQLQKFPSRVTDNSAEPK-----DVQETVFSV 556  
 Db 427 -ALPTDQAQTNEWAVINPQTFPQPKAKNTFHNNSTKTSAFNHSRVPNPDLPET---- 480  
 Qy 557 DVGNVNVGRDGPBVFTAEAVRPSSENFTNDYAIRAGVSESSVDETVMSVADMK--- 613  
 Db 481 -----PTK--ETSEYPNTSPPMWSSGS1KSEVPAEARNRG 516  
 Qy 614 ---EASVKTLAG----- 623  
 Db 517 NOHLINGSAKASAASGHNOKRRRKPTSASRERIPDNRPHSSRRRRTPAMTIEGKRLV 576  
 Qy 624 --VAIGLISL-----FSQYFLKSSSFORKDKDMVSMEDVATIGSVRADSB 669  
 Db 577 WRVPISLVSLVFWVLATTCRWLKNLPPQSPPDQLQFVQINQPLPDPNPKRE 636  
 Qy 670 ALPRMDARTAENIVSKQOKIKSLAFGDPHR-EMPLPVLDGMLKIKWTDRAETAQIQLGVY 729  
 Db 637 EGPLTNAB-ASBVHTWLSTKAALGPNHEKNNQEQILTSALSQPR-LIAQONKUDNR 694  
 Qy 730 --DYTILKLSVDSVTSVSDGTRALVEATLBESACSLDVLVHPENNATDVRTYTYTRVFW 786  
 Db 695 RKEDHSLSKLESVEKIGLFD--RAAVERATVKVTOVLYENNOPKNSND--KLRVRLDLR 750  
 Qy 787 SRSQWKITEGSTL 799  
 Db 751 ERGKWRQIOTSvV 763

RESULT 15  
 US-10-424-599-177901  
 ; Sequence 177901, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovacic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(5323)B  
 ; CURRENT APPLICATION NUMBER: US/10/424, 599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 28564  
 ; SEQ ID NO: 177901  
 ; LENGTH: 204  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRP3847\_131660C.1.pep  
 ; US-10-424-599-177901

Query Match Best Local Similarity 51.2%; Score 499; DB 15; Length 204;  
 ; Matches 104; Conservative 39; Mismatches 58; Indels 2; Gaps 2;  
 Qy 601 ETIVMSVADMKAEASVKKLAVGATVGLLISRSQKVF-LKSSSFORKDKDMVSMEDVAT 659  
 Db 2 ETRKGWVITEEJKHASVQMCAGWVIGLTVUGLKLPTGSPSPRKWMCSAMVSTDIN 61  
 Qy 660 IGSVRADS-EALPRMDARTAENIVSKQOKIKSLAFGDPHR-EMPLPVLDGMLKIKWTD 718  
 Db 62 LGSLGDBKEVKQFQMDARVABALVVKWQSVSEAGPDCIGRHLHEVLDGEMKWTDR 121  
 Qy 719 AAEATAQOLGLVYDVTLLKLSVSDVTSVSDGTRALVEATLBESACSLDVLVHPENNATDVRTY 778  
 Db 122 ARTAERGWSDYDLENDLUDSUTSINGRGAVENTILKESTHINAVGPHONDASNSR 181  
 Qy 779 TRYEVFMWSKWKITESVULAS 801  
 Db 182 TRYEMSFPTGAEWKIVEGAVIDES 204

/ Fri Jun 10 09:57:55 2005

Search completed: June 10, 2005, 01:50:42  
Job time : 477 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

Om nucleic - nucleic search, using sw model

Run on: June 9, 2005, 15:02:56 ; Search time 11188 Seconds  
(without alignments)  
12476.017 Million cell updates/sec

Title: US-10-600-070-3

Perfect score: 3667  
Sequence: 1 tggtctgcattaaggaaat.....ctataacattaaagggtacaa 3667

Scoring table: IDENTITY\_NUC  
Gapext 10.0 , Gapext 1.0

Searched: 34239544 seqb, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_htc1:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gbseq:\*

9: gb\_gbseq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

#### ALIGNMENTS

RESULT	REFERENCE	TITLE	KEYWORDS	SOURCE	ORGANISM	DESCRIPTION
1	CNS09YJH	CNS09YJH			Arabidopsis thaliana	919 bp mRNA linear HTC 04-FEB-2004
2	BX841670				Arabidopsis thaliana	Full-length cDNA Complete sequence from clone CNS09YJH of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).
3	AJ1988415	CNS09YJN				
4	BX841670	GI:42406830				
5	BX841670.1					
6						
7						
8						
9						

REFERENCE	AUTHORS	TITLE	COMMENT
1	(bases 1 to 919)		The Sequences are based on single pass reads.
			Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
			Genoscope members carried out sequencing and annotation : Castelli V., Aury J.-M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M., URGV INRA : Clepet C., Caboche M.
			Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences). 5 Prime and 3 prime are assembled with Phrap. http://www.genoscope.cnrs.fr/extreme/sequences/Banque_Projet_EP/Full_length

FEATURES	REFERENCE	TITLE	COMMENT
Source	Location/Qualifiers		
	1. -919		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/strain="Col-0"		
	/db_xref="taxon:3702"		

1	2307	9	CL965374
2	212.6	5.8	CN94655
3	21.0	5.7	AL952467
4	208.8	5.7	BL268376
5	5.3	2307	NP19C111
6	239.6	6.5	AV812946
7	237.2	6.5	AV812946
8	293.4	8.0	CO079829
9	265.4	7.2	CN94655
10	254	6.9	CR46014
11	254	6.9	Medicago
12	252.6	6.9	CO117047
13	252	6.9	GR_BD019
14	252	6.9	CO075595
15	252	6.9	GR_Ba30N
16	252	6.9	CO071968
17	252	6.9	GR_Ba30N
18	239.6	8.1	BX046755
19	237.2	8.0	PP_LBa002
20	212.6	8.0	CN94655
21	212.6	5.8	CN94655
22	21.0	5.7	CN94655
23	208.8	5.7	AL952467
24	5.3	2307	OB_Ba007

gene	ORIGIN	/clone="GS11LS82ZC08" /tissue type="Adult vegetative tissue" /plasmid="pCMVSPORT_5" /gene="At5g42480"
Qy	Query Match Best Local Similarity 19.8%; Score 725.2; DB 3; length 919; Matches 916; Conservative 0; Mismatches 3; Indels 169; Gaps 2;	
Qy	469 ATAGACACTCCCTTGAAGACTCTCTAGTCAGTCAGTCGGCATTTGTTCTCCCATTCATA 528 1 AAGACACTCCCTTGAAGACTCTCTAGTCAGTCAGTCGGCATTTGTTCTCCCATTCATA 60	
Db	529 TCGGATTAACACGGGAGCAAGTCAGTCAGTCGGCATTTGTTCTCCCATTCATA 588 1 TCGGATTAACACGGGAGCAAGTCAGTCAGTCGGCATTTGTTCTCCCATTCATA 60	
Db	61 TCGGATTAACACGGGAGCAAGTCAGTCAGTCGGCATTTGTTCTCCCATTCATA 120	
Qy	589 TGTCGCCGAGAATGGCCGAGCTCTCCGACTCAATTGCGATTCATTGCGATTC 648	
Db	121 TGTCCCGCAGAAATGCGGAGCGCTCTCCGACTCAATTGCGATTC 180	
Qy	649 TCCCTCTCTCTCTGCGCACCGCACCCACCCACTCTCTCTCTGCACCATCT 708 181 TCCTCTCTCTGCGCACCGCACCCACCCACTCTCTCTGCACCATCT 240	
Db	709 ATTTGATCTCCCTTGCCACCGCACCCATCCCTGATTCACCAGTAGGAGCT 768	
Qy	241 ATTTGATCTCCCTTGCCACCGCACCCATCCCTGATTCACCAGTAGGAGCT 300	
Db	769 CAACACATTCTAACCATGATCAGAGGACATGAGCTAGAGCTAGGGTTGAAACG 828	
Qy	301 CAACACATTCTAACCATGATCAGAGGACATGAGCTAGAGCTAGGGTTGAAACG 360	
Qy	829 CCCAACATTGGTTTCAAGGACGAGGCTTAATGCGCGAGAAGAGATCTTCAAGCT 888 361 CCCAACATTGGTTTCAAGGACGAGGCTTAATGCGCGAGAAGAGATCTTCAAGCT 420	
Db	889 TSGAACACTCTGCTAAATCTCGCTTGAAGAGAGTCAATGAGCTTCTGTGAT 948	
Qy	421 TSGAACACTCTGCTAAATCTCGCTTGAAGAGAGTCAATGAGCTTCTGTGAT 480	
Db	949 GAGAGAGCTACGCTACGTGTTCTGGATAAGGAAATTGGAAATTGGAA 1008	
Qy	481 GAGAGAGCTACGCTACGTGTTCTGGATAAGGAAATTGGAAATTGGAA 516	
Qy	1009 TAAGTTCTCGTTTAATTGATGAAATGGAAGAACATTATCTAGTGA 1068 517 -----AA 518	
Db	1069 GGTCTCTGGGCTCTGTGATGCAAGAAGSTGGTGTGAGCTGAGAPAGTTCTTCGGT 1128 519 GGTCTCTGGTGTCTCTGTTGAGTGTGAGAAGGGTGTGAGTGTCTTCGGT 578	
Db	1129 TCTGGAGCTCTTAAGGAGGGTCTCTAGTCGTTAACGAGTGTGTTAGT 1188 579 TCTGGAGCTCTTAAGGAGGGTCTCTAGTCGTTAACGAGTGTGTTAGT 638	
Qy	1189 TATGGCGTTGCGTTCTCGATGTCCTCGAGGGTGTAGGCGATTCAGCTGATT 1248 639 TATGGCGTTGCGTTCTCGATGTCCTCGAGGGTGTAGGCGATTCAGCTGATT 698	
Qy	1249 TATTAAGCTGTTAGAGTTGTTAGGAAGCTTGAAGCTTTCAGGTAGTTGACTTGC 1308 699 TATTAAGCTGTTAGAGTTGTTAGGAAGCTTGAAGCTTGC 742	
Db	1309 TTGGTAATTGAGGAGGAGCTGCTTATAAGACTTCTGATTGACTTGTATT 1368 743 -----AA 742	
Qy	1369 GAGCTCTGTTAGGAGGAGGAGCTGCTTACAATG 1428	
Db	743 -----AGGAGGAGGAGCAGTGAATTGCGATCCGATTAGTCACAATTGA 791	
Qy	1429 TGGACTTTGGAGAGATACCTCGCGTTATGCTGAGCTRACTTGCTTACCGCTTG 1488	
Db	792 TGAACITGGAGAGATACCTCGCGTTATGCTGCTTGGCTRACTTGCTTACCGCTTG 851	
Qy	1489 TGGATTTAGCTGGAAGAGCTAATGTTTAACGGTTGGGATATTGTTGGCT 1548 852 TGTGATTTAGCTGGAAGAGCTAATGTTAAGGGTGGGGATATTGTTGGCT 911	
Db	1549 TGTGGAG 1556	
Db	912 TGTGGAG 919	
<b>RESULT 2</b>		
CNS02ZSS		
DEFINITION	CNS02ZSS	mRNA
LOCUS	CNS02ZSS	linear
VERSION	BX833051.1	HTC
KEYWORDS	HTC; GS11LS82ZC08	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
JOURNAL	Arabidopsis thaliana (thale cress)	
REFERENCE		
AUTHORS	Castelli, V., Aury, J.-M., Jaijou, O., Wincker, P., Clepet, C., Menard, M., Crouaud, C., Quétier, F., Scarpelli, C., Schachter, V., Bukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
COMMENT	1 (bases 1 to 741)	
ARTICLE	Unpublished	
TITLE	2 (bases 1 to 741)	
JOURNAL	Genoscope - Centre National de Séquençage : Submitted (18-NOV-2003) Genoscope - Centre National de Séquençage : BP 191 91006 EVRY Cedex - FRANCE (E-mail : seqrof@genoscope.cnrs.fr)	
COMMENT	- Web : <a href="http://www.genoscope.cnrs.fr">www.genoscope.cnrs.fr</a>	
ARTICLE	The sequences are based on single pass reads.	
LIBRARY	Life Technologies (a division of Invitrogen) members carried out full-length librairies construction : Temple G.	
GENOME	Genoscope members carried out sequencing and annotation : Castelli V., Aury J.-M., Jaijou O., Wincker P., Menard M., Crouaud C., Schachter V., Weissenbach J., Salanoubat M.	
URGV INRA	URGV INRA : Clepet C., Caboche M.	
ANNOTATION	Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with phrap, http://www.genoscope.cnrs.fr/extreme/sequences/Banque_Projet_EPF_Full	
LENGTH	length	
FEATURES	<a href="http://www.genoscope.cnrs.fr/cgi-bin/ggb/gb?source=Arabidopsis">http://www.genoscope.cnrs.fr/cgi-bin/ggb/gb?source=Arabidopsis</a> . Location,Qualifiers	
source	1. .741	
	/organism="Arabidopsis thaliana"	
	/mol_type="mRNA"	
	/strain="Col-0"	
	/db_xref="taxon:3702"	
	/clone="GS11LS82ZC08"	
	/tissue type="Silique"	
	/plasmid="pCMVSPORT_6"	
	/gene="At5g42480"	
ORIGIN		
Qy	Query Match Best Local Similarity 11.9%; Score 436.2; DB 3; Length 741; Matches 548; Conservative 0; Mismatches 18; Indels 91; Gaps 1;	

QY	COMMENT
Db	2970 TTCTTCTGTGATTTCCTTCGATTAGGGCATCTAGAGATCGAGATCTGGAG 3029
Db	176 TTCTTCTATGGATCTGTCCTACATGGCTAGTCAGCTGAGATCAGAG 235
Qy	3030 CACTTCCAGAATGGATCTAGACTTGAGAGATATAGTATCCAACTGGAGATTA 3089
Db	296 AGTCTCTGCTTGGCTGATCACCGCTAGAACTGGAGATTA 295
Db	236 CACTTCCAGAATGGATCTAGACTTGAGATTA 295
Qy	3150 TCTAACATTCAATCAATTGTGTGAACACTGTGGTGCATGATTAGTGTCCTCT 3209
Db	341 ----- 340
Db	3210 TGATTCTGTATTGTAGTTGATGGCTGAGGGATAGTGGACTGAGACCA 3269
Qy	385 GCTGAACATGCGAGCTGGGGTGTGTTATGATTA 3444
Db	341 ----- -AGCTTGGATGGCTGAGATTA 340
Qy	3270 GCTGAACATGCGAGCTGGGGTGTGTTATGATTA 3329
Db	385 GCTGAACATGCGAGCTGGGGTGTGTTATGATTA 3329
Qy	3330 AGTGTGACAGCTGAGCTGGGGTGTGTTATGATTA 3389
Db	445 AGTGTGACAGCTGAGCTGGGGTGTGTTATGATTA 3389
Qy	3390 GCTGTGCTATGATGGCTGAGCTGGGGTGTGTTATGATTA 3449
Db	505 GCTGTGCTATGATGGCTGAGCTGGGGTGTGTTATGATTA 3449
Qy	3450 ACAGAGATGAGTTCTGGTCAAGTCAGAAACATGCTGAGACCTACCA 3509
Db	565 ACAGAGATGAGTTCTGGTCAAGTCAGAAACATGCTGAGACCTACCA 3509
Qy	3510 GCATCATATACTCATATGATGGCTGAGCTGGGGTGTGTTATGATTA 3569
Db	625 GCATCATATACTCATATGATGGCTGAGCTGGGGTGTGTTATGATTA 684
Qy	3570 ATCTCTCTAGTAGTGTGATGGCTGAGCTGGGGTGTGTTATGATTA 3625
Db	685 ATCTCTCTAGTAGTGTGATGGCTGAGCTGGGGTGTGTTATGATTA 741
RESULT 3	
CNS097NM	CNS097NM 741 bp mRNA linear HRC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUTSIL522A05 of silique of strain col-0 of Arabidopsis thaliana (thale cress).
DEFINITION	
ACCESSION	BX833489.1 GT:42455179
VERSION	
KEYWORDS	HTC; GSLT; cDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana Bukaryota; Viridiplantea; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsidae.
REFERENCE	1 (bases 1 to 741)
AUTHORS	Carroll, V., Aury, J.M., Jaillet, O., Wincker, P., Clepet, C., Menard, M., Crumaud, C., Quétier, F., Scarpelli, C., Schachter, V., Tempé, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
TITLE	Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL	Unpublished 2 (bases 1 to 741) Genoscope.
REFERENCE	Direct Submission
AUTHORS	Submitted (18-NOV-2003) Genoscope - Centre National de Séquençage : BP 191-91006 EURY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
FEATURES	source
ORIGIN	
Query	Match 11.9%; Score 434.6; DB 3; Length 741;
Query	Best Local Similarity 83.3%; Pred. No. 8; 7e-101; Matches 547; Conservative 0; Mismatches 19; Indels 91; Gaps 1;
Query	2970 TTCTTCTATGGATCTGATGGCTACCATAGGGCTAGTCAGAGCTGAG 3029
Db	176 TTCTTCTATGGATCTGATGGCTACCATAGGGCTAGTCAGAGCTGAG 235
Qy	3030 CACTTCCAGAATGGATCTAGACTGGAGAAATAGTATGATTA 3089
Db	236 CACTTCCAGAATGGATCTAGACTGGAGAAATAGTATGATTA 295
Qy	3150 GCATCATATACTCATATGATGGCTGAGCTGGGGTGTGTTATGATTA 3569
Db	625 GCATCATATACTCATATGATGGCTGAGCTGGGGTGTGTTATGATTA 684
Qy	3210 TGATTCTGTATTGTAGTTGATGGCTGAGGGATAGTGGACTGAGACCA 3269
Db	341 ----- -AGCTTGGATGGCTGAGATTA 340
Qy	3270 GCTGAACATGCGAGCTGGGGTGTGTTATGATTA 3329
Db	385 GCTGAACATGCGAGCTGGGGTGTGTTATGATTA 3329
Qy	3330 AGTGTGACAGCTGAGCTGGGGTGTGTTATGATTA 3389
Db	445 AGTGTGACAGCTGAGCTGGGGTGTGTTATGATTA 3389
Qy	3390 GCTGTGCTATGATGGCTGAGCTGGGGTGTGTTATGATTA 3449
Db	505 GCTGTGCTATGATGGCTGAGCTGGGGTGTGTTATGATTA 3449
Qy	3450 ACAGAGATGAGTTCTGGTCAAGTCAGGGTGAATCTGAGCTGCTGTT 3509
Db	565 ACAGAGATGAGTTCTGGTCAAGTCAGGGTGAATCTGAGCTGCTGTT 3509
Qy	3510 GCATCATATACTCATATGATGGCTGAGCTGGGGTGTGTTATGATTA 3569
Db	625 GCATCATATACTCATATGATGGCTGAGCTGGGGTGTGTTATGATTA 684
Qy	3570 ATCTCTCTAGTAGTGTGATGGCTGAGCTGGGGTGTGTTATGATTA 3625
Db	685 ATCTCTCTAGTAGTGTGATGGCTGAGCTGGGGTGTGTTATGATTA 741

RESULT 4  
AI98415/C  
LOCUS AI98415 561 bp mRNA linear EST 08-SEP-1999  
DEFINITION thaliana cDNA clone 701545606, mRNA sequence.  
VERSION AI98415  
KEYWORDS EST  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
AUTHORS 1 (bases 1 to 561)  
Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,  
Wang,X., Hillman,J., Gugler,K., Kim,C., Doyle,M., Brzak,P.,  
Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D.,  
Tan,R., Rose,M., Warren,B., Ion,B., Kasbury,K., Borillo,C.,  
Carrio,T., Policky,J., Suzuki,G., Argentine,C., Shah,S., and  
Nobriga,A., Murry,L., Turner,C., Krikorian,S., Elder,L., and  
Hanson,D.  
TITLE Arabidopsis thaliana Gene Expression MicroArray  
JOURNAL Unpublished (1999)  
COMMENT Contact: David Smoller, Ph.D.  
Genome Systems, Inc., a wholly owned subsidiary of Incyte  
Pharmaceuticals, Inc.  
4633 World Parkway Circle, St. Louis, MO 63134, USA  
Tel: 877-577-2733  
Fax: 314-427-3324  
Email: service@genesystems.com.

FEATURES source  
1. . 561 Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Col-0"  
/db\_xref="taxon:3702"  
/clone="701545606"  
/tissue\_type="rosette"  
/dev\_stage="7 weeks"  
/clone\_lname="A. thaliana, Columbia Col-0, rosette-2"  
/note="vector: pSPORE; Site 1: NotI; Site 2: Sall; cDNA  
library was derived from untreated rosette tissue from  
Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks.  
Plants were grown in 1:1:1 peat moss/vermiculite/perlite  
soil at 22 deg. C +/- 3 deg. C under constant light, and  
watered with fertilizer. cDNA synthesis was initiated  
using a Not-I-oligo(dT) primer. Double-stranded cDNA was  
blunted, ligated to Sall adaptors, digested with NotI,  
size-selected, and cloned into the NotI and Sall sites of  
the pSPORE vector."

ORIGIN

RESULT 5  
B2437564/C  
LOCUS B2437564 785 bp DNA linear GSS 13-DEC-2002  
DEFINITION BONRN72R BO 1.6-2 KB tot Brassica oleracea genomic clone BONRN72,  
genomic survey sequence.  
VERSION B2437564  
VERSION B2437564.1 GI:26691135  
KEYWORDS GSS  
SOURCE Organism  
Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosid II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 785)  
REFERENCE Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
AUTHORS  
TITLE Unpublished (2001)  
JOURNAL Other\_GSSB: BONRN72F  
COMMENT Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0209  
Email: ctown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.  
FEATURES source  
1. . 785 Location/Qualifiers  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="R01000DH3"  
/db\_xref="taxon:3712"  
/clone="BONRN72"  
/clone\_lname="BO 1.6-2 KB tot"  
/note="vector: phosI; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into phosI using BstXI linkers"  
ORIGIN

Query Match 9.9%; Score 363.6; DB 8; Length 785;  
Best Local Similarity 73.3%; Pred. No. 2.1e-82; Matches 582; Conservative 0; Mismatches 109; Indels 103; Gaps 5;

Db 3210 TGATCTGTATTATAGGTTTGGATGGCGATGCTGAAGATTCGACTCACAGAGCA 3269  
Db 369 ----- AGGTTTGGGGGAATGCTGAGATTGGACAGACAGA 326  
Qy 3270 GCTGAAACTGCCAGCTGGTTGTTATGATTAACATCTTGAAACTATCTGTGAC 3329  
Db 325 GTGAAACTGCCAGCTGGTTGTTATGATTAACATCTTGAAACTATCTGTGAC 266  
Qy 3330 AGTGTGAGAGCTCAGATGAGAACCCGCTCTGGGAGAAGCACTGGAGAGCT 3389  
Db 265 AGTGTGAGAGCTCAGATGAGAACCCGCTCTGGGAGAAGCACTGGAGAGCT 206  
Qy 3390 GCTGTGCTATCGATTGGTCATCCGAGAACATCTACTGATGTCAGAACCTACA 3449  
Db 205 GCTGTGCTATCGATTGGTCATCCGAGAACATCTACTGATGTCAGAACCTACA 146  
Qy 3450 ACAGATACTGAAAGTTCTGGTCCAACTCAGGGTGAAGAAATCACTGAGGCTGTGCT 3509  
Db 145 ACAGATACTGAAAGTTCTGGTCCAACTCAGGGTGAAGAAATCACTGAGGCTGTGCT 86  
Qy 3510 GCATCATATACTCATGAGCTGAGCTGAGCTGAGGAGATCTCTGTGCT 3569  
Db 85 GCATCATATACTCATGAGCTGAGCTGAGGAGATCTCTGTGCT 26  
Qy 3570 ATCTCTCTCTAGTTAGTTAT 3594  
Db 25 ATCTCTCTCTAGTTAGTTAT 1

Query Match 9.9%; Score 363.6; DB 8; Length 785;  
Best Local Similarity 73.3%; Pred. No. 2.1e-82; Matches 582; Conservative 0; Mismatches 109; Indels 103; Gaps 5;

Db 3150 TCTPACAATTCAATCAATGTCGAAACTGTTGACATGATPATAGTCGTCGCCTGTT 3209  
Db 369 ----- 370



AUTHORS	Roose, M.L., Ye, X., Federici, C.F., Close, T.J., Fenton, R.D., Wanamaker, S., Choi, Y. and Kingan, T.
TITLE	Development of EST Resources and New Genetic Markers for California Citrus - Poncirus trifoliata CTV-challenged phloem - UCR Unpublished (2003)
JOURNAL	Department of Botany & Plant Sciences, University of California Riverside, CA, 92521-0124, USA
COMMENT	Contact: Mikeal Roose Tel: 909/774137 Fax: 909/774437 Email: mikeal.roose@ucr.edu
FEATURES	Seq primer: T3;
SOURCE	location/Qualifiers 1. .897 /organism="Poncirus trifoliata" /mol_type="mRNA" /cultivar="Pomeyoy OP" /db_xref="Taxon:37690" /clone="UCRPT01_01_F12" /tissue_type="Phloem" /dev_stages="10 - 30 cm shoots" /lab_host="E. coli TUC12"; /clone_id="Poncirus trifoliata CTV-challenged cDNA library - UCR" /notes="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: Ecoli; Site_2: Axio; Plants were grown in the greenhouse at University of California, Riverside. The scion was a open-pollinated (very probably selfed) seedling of Poncirus trifoliata cv Pomeyoy that was selected as homozygous for the Ctv resistance gene. The rootstock was sweet orange infected with citrus tristeza virus (CTV) isolate T514 over 1 year before sampling (CTV infects sweet orange, but not genotypes carrying the Ctv resistance gene). Shoots 10-30 cm long were harvested in October 2000, and the green phloem (bark) was removed and frozen quickly in dry ice. Total RNA was extracted using Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA library was made, and 0.5 million primary lambda cDNA clones were in vivo excised to give a population of pBluescript SK(-) phagemids. All steps to this point were performed in the MU Roos lab at the University of California, Riverside by X. Ye. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequence determined bi-directionally using an ABI3730 at the University of California Riverside Genomics Institute, Core Instrumentation Facility. (Choi, Kingan). Chromatogram files were transmitted to UC Riverside (by Choi), then processed at UC Riverside (by Wanamaker) using the HARVEST pipeline ( <a href="http://harvest.ucr.edu">http://harvest.ucr.edu</a> ) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."
ORIGIN	Query Match 9.2%; Score 336.4; DB 6; Length 897; Best Local Similarity 63.9%; Pred. No. 2.4e-75; Matches 574; Conservative 0; Mismatches 241; Indels 83; Gaps 1;
QY	1391 GCAAGTAGCCCTGCCACCGGATTAGTGACAAATTGATGAGGTTGGAGAGGATCACT 1450
Db	80 GCGAGTAGTCCTCCAGATTCGAGCAGATGAGGATCACT 139
QY	1451 CGCGGTAGTGTCTGGAGCTACTTGCTTACCCCTGGTGTGATTAGCTGGAGAAGA 1510
Db	140 CGCGGTGTGTGGAGCTTCTCCCTAGCTTACCCCTGGTGTGAGTGTCAAGCAAGAGA 199
QY	1511 CTAAATGGTTAGGGGTTGGGAATATTGCTGCTGAGGGGGAGCATCA 1570
Db	200 GAGGAAAGACTTGTATGGCTAACATACTATGGCTGGAGGGGGAGCACT 259
QY	1571 GCTCTGTGGGGTTGACCCGTGAGAGTTATGATGAGGCTTGTAGGATGACA 1630
RESULT	8
BQ834167	534 bp mRNA linear EST 14-MAR-2003
LOCUS	BQ834167
DEFINITION	ALEST0128 Arabidopsis lyrata Inflorescence PCR Library
ACCESSION	BQ834167
VERSION	BQ834167.1 GI:28951482
KEYWORDS	EST.
SOURCE	Arabidopsis lyrata
ORGANISM	Arabidopsis lyrata
REFERENCE	Eukaryota; Vidioplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	Barrie, M., Bustamante, C.D., Yu, J. and Purugganan, M.D.
TITLE	Selection on rapidly evolving proteins in the <i>Arabidopsis</i> genome
JOURNAL	Genetics 163 (2), 723-733 (2003)
PUBLISHED	2005/04/05
COMMENT	Contact: Barrie M Department of Genetics North Carolina State University 3513 Gardner Hall, Box 7614, Raleigh, NC 27695, USA Tel: 919 515 1761 Fax: 919 515 1695 Email: mbarrie@unity.ncsu.edu

FEATURES	REFERENCE
KEYWORD	AUTHORS
ORGANISM	robida; eurosis II; Brassicales; Brassicaceae; Arabidopsis /organism="Arabidopsis lyrata" /mol_type="mRNA" /cultivar="Karuhamaki" /db_xref="taxon:59689" /clone="PWB1-D03" /tissue_type="Inflorescence" /clone_Type="Arabidopsis lyrata Inflorescence pCMV-PCR /note="Vector: pCMV-PCR (Stratagene); Created using PCR Library Construction kit (Stratagene)"
ACCESSION	ORIGIN
RESULT 9	Query Match 8.6%; Score 315.8; DB 5; Length 534; Best Local Similarity 76.9%; Pred. No. 4.4e-70; Mismatches 42; Indels 91; Gaps 1; Matches 442; Conservative 0; Mismatches 51
AV830764	TITLE JOURNAL
LOCUS	/unpublished (2002) Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: msek@riic.riken.go.jp
VERSION	COMMENT
DEFINITION mRNA sequence.	An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XbaI was ligated to modified lambda PLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and Sall. This clone is in a modified pBlue-script vector. Please visit our web site ( <a href="http://www.gsc.riken.go.jp/e/plant/index_e.html">http://www.gsc.riken.go.jp/e/plant/index_e.html</a> ) for further details.
KEYWORD	FEATURES source
EST	1. 320 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="RAFL09-76-G11" /dev_stages="plants at various developmental stages from germination to mature seeds" /lab_host="DH10B" /clone_Lib="RAFL09" /note="Site 1: BamHI; Site 2: Sall; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
ORGANISM	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
AV830764	Matches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
LOCUS	FEATURES source
VERSION	1. 320 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="RAFL09-76-G11" /dev_stages="plants at various developmental stages from germination to mature seeds" /lab_host="DH10B" /clone_Lib="RAFL09" /note="Site 1: BamHI; Site 2: Sall; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
DEFINITION mRNA sequence.	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
KEYWORD	Matches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
EST	1. 320 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="RAFL09-76-G11" /dev_stages="plants at various developmental stages from germination to mature seeds" /lab_host="DH10B" /clone_Lib="RAFL09" /note="Site 1: BamHI; Site 2: Sall; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
ORGANISM	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
ACCESSION	RESULT 10
RESULT 9	REFERENCE
AV830764	1. (bases 1 to 320) Seki, M., Narusawa, M., Ishida, J., Kamiya, A., Satoh, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibusawa, R., Shinagawa, A., Muramatsu, M., Hayashizaki, Y., and Shinozaki, K.
LOCUS	Large scale analysis of Arabidopsis full-length cDNA (2002b)
VERSION	Unpublished (2002)
DEFINITION	Unpublished (2002)
KEYWORD	Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: msek@riic.riken.go.jp
EST	An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XbaI was ligated to modified lambda PLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and Sall. This clone is in a modified pBlue-script vector. Please visit our web site ( <a href="http://www.gsc.riken.go.jp/e/plant/index_e.html">http://www.gsc.riken.go.jp/e/plant/index_e.html</a> ) for further details.
ORGANISM	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
ACCESSION	FEATURES source
RESULT 9	1. 320 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="RAFL09-76-G11" /dev_stages="plants at various developmental stages from germination to mature seeds" /lab_host="DH10B" /clone_Lib="RAFL09" /note="Site 1: BamHI; Site 2: Sall; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
AV830764	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
LOCUS	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
VERSION	1. 320 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="RAFL09-76-G11" /dev_stages="plants at various developmental stages from germination to mature seeds" /lab_host="DH10B" /clone_Lib="RAFL09" /note="Site 1: BamHI; Site 2: Sall; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
DEFINITION	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
KEYWORD	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
EST	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
ORGANISM	Query Match 8.5%; Score 311.8; DB 1; Length 320; Best Local Similarity 98.4%; Pred. No. 4.1e-69; Mismatches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
ACCESSION	RESULT 10
RESULT 9	REFERENCE
AV830764	1. (bases 1 to 320) Seki, M., Narusawa, M., Ishida, J., Kamiya, A., Satoh, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibusawa, R., Shinagawa, A., Muramatsu, M., Hayashizaki, Y., and Shinozaki, K.
LOCUS	BZ765587
VERSION	320 bp DNA linear GSS 13-MAR-2003
DEFINITION	BZ765587
KEYWORD	SALK_131837.28-70.x Arabidopsis thaliana TDNA insertion lines
EST	SALK_131837.28-70.x Arabidopsis thaliana genomic clone SALK_131837.28-70.x, genomic survey sequence.
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

VERSION	BZ765587.1	GI	28938140	DEFINITION	MBERKH28TRB mth2 Medicago truncatula genomic clone 73FB, genomic
KEYWORDS	GSS.			ACCESSION	Survey sequence.
SOURCE	Arabidopsis thaliana (thale cress)			VERSION	CG961431
ORGANISM				VERSION	CG961431.1
REFERENCE	1 (bases 1 to 320)			KEYWORDS	GI:39883077
AUTHORS	Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Prednis, L.,			ORGANISM	GS.
	Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Shinn, P., Zimmerman, J. and Ecker, J.R.			REFERENCE	Medicago truncatula (barrel medic)
TITLE	A Sequence-Indexed Library of Insertion Mutations in the			ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.
JOURNAL	Arabidopsis Genome			REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
COMMENT	Unpublished (2001)			ORGANISM	1 (bases 1 to 954)
CONTACT	Joseph R. Ecker			AUTHORS	Town, C.D., Shetty, J., Koo, H. and Feldblum, T.F.
SALK INSTITUTE GENOMIC ANALYSIS LABORATORY (SIGNAL)	The Salk Institute for Biological Studies			TITLE	Sequencing of BAC ends from <i>Medicago truncatula</i>
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA				JOURNAL	Unpublished (2003)
Tel: 858 453 4100 x1752				COMMENT	Other GSes
Fax: 858 558 6379					Contact: Chris Town
Email: ecker@salk.edu					TIGR
This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of AT5g12480.					9712 Medical Center Drive, Rockville, MD 20850, USA.
Class: TDNA tagged.					Tel: 301-835-3523
FEATURES	TDNA tagged.				Fax: 301-838-0208
SOURCE	Location/Qualifiers				Email: catow@tigr.org
	1. .320				Seq primer: CAGGAACGCTATGACC
/organism="Arabidopsis thaliana"					Class: BAC ends
/mol_type="genomic DNA"					Location/qualifiers
/ecotype="Col-0"					1. .954
/db_xref="TAXON:3702"					/organism="Medicago truncatula"
/clone="SALK_13837_28.70_X"					/mol_type="genomic DNA"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"					/cultivar="genotype Al7"
/note="PCR was performed on <i>Arabidopsis thaliana</i> lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at <a href="http://signal.salk.edu/tDNA_protocols.html">http://signal.salk.edu/tDNA_protocols.html</a> "					/db_xref="TAXON:3880"
ORIGIN					/clone="73FB8"
Query Match	8.4%; Score 307.2; DB 8; Length 320;				/clone_lib="mth2"
Best Local Similarity	97.5%; Pred. No. 6.4e-68;				/note="Vector: pBELoBAC1; Site_1: HindIII; Site_2: HindIII; Cook, D.R. and Kim, D.J., unpublished"
Matches	312; Conservative 0; Mismatches 8; Indels 0; Gaps 0;				
QY	343 TTTCAGTAGCTGTGATGTTTGATTAATATACTCAAAATCAATTCCAT	402			
Db	1 TTTCAGTAGCTGTGATGTTTGATTAATATACTCAAAATCAATTCCAT	60			
QY	403 AACCCCTAGACGCCAACAGCTCTCTCATATGTAACAGAACAAAGTTTGAGTA	462			
Db	61 AACCCCTAGACGCCAACAGCTCTCTCATATGTAACAGAACAAAGTTTGAGTA	120			
QY	463 GCCTAAAGAACCTCCATGGAGCTGAGCACGGTGGCTTGGCTCTGCCATTC	522			
Db	121 GCCTAAAGAACCTCCATGGAGCTGAGCACGGTGGCTTGGCTCTGCCATTC	180			
QY	523 CAAATTATGCCGATACACCCGGGAGCAGAACACTCCAGCTGAGCCACAAACCTCTCA	582			
Db	181 CGTATATGCCGATACACCCGGGAGCAGAACACTCCAGCTGAGCCACAAACCTCTCA	240			
QY	583 ACTATCTCCGCCASCAAATGGGCCACCGCTTCTCCGACTCAATTCACTTC	642			
Db	241 ACTATCTCCGCCAGCAATGGCCACCGCTCTCCACTCAATTCACTTC	300			
QY	643 GATTCCTCTCCCTCTCTCT 662				
Db	301 GATTCCTCTCCCTCTCT 320				
RESULT	11				
LOCUS	CG961431/c				
	954 bp	DNA	linear	GSS	15-DEC-2003
QY	1812 TTATTCGTTATCTCCATCTCTTTAGGTGACHTTGTAGCAACACTAG	437			
Db	1752 CAAATTCTCCGCCCTCTAGTTATGAGTTTACATTATCTCTCA	497			
QY	436 CAAATTCCAGTGAAGTTTGAGCTCTATGGGTCAGTCAGTGGTGCAG	377			

ORIGIN

Query Match	8.1%	Score	295.6;	DB	5;	Length	631;
Best Local Similarity	71.7%	Pred.	No.	7.8e-65;			
Matches	407;	Conservative	0;	Mismatches	149;	Indels	12;
						Gaps	1;

Db 376 CTTGTGAGTGGTAAAGGCCACATCTTATCCAGAGTGCTGATAATTTCATCACTTCA 317

Qy 1872 GCAGGCTAAGTAATGGCTATGGAGATTCCTGCGATGTTGATGATAC-----ACGGAA 1925

Db 316 ACDAACTAAGTTACAATGAACTCTCCCTCTGTTACACTCCATGAGATGGA 257

Qy 1926 TAATGGGAGATAGACTTGGCTAGAACAGGGACTCTTGACTCTGCTTATAGCAAGT 1985

Db 256 GAGAGAGAGTTGATTGCAATGAGAAGGGCTGCTGTTGACTGCTGTTGGGAGCT 197

Qy 1986 TGAATGCCGATGTCGTTGAGAACTGAGCTGTTGAGAAGGAGCTGAGATCCAG 2044

Db 196 TGATCAATTCGATCATGGTGGCTCTGATGAGACTCTCACCTATGAAACCAT 137

Qy 2045 CTATGGGAGTTGTTGGAGAACTGAGCTGCTGCTGCTTC-CTAGTTGAGACAC 2103

Db 136 CTATTTAGACTTATATGGAAACGAAAGTGTGAGAAGTGTGAGATCTTCCTGA 77

Qy 2104 CTATGCAATGTGTTGGAAACCTGGCAGGGGTTGCTGCTTC-CTAGTTGAGACAC 2162

Db 76 CTGTGAAATGTGTTGGAGACATGCTGAGTGAAGGTTGTTCTCTAGTTAGATAC 17

Qy 2163 CATAGATAAAATT 2178

Db 16 TAAGAGACAACTT 1

RESULT 12

LOCUS BU046755 631 bp mRNA linear EST 26-AUG-2002

DEFINITION PP\_LBA0027104f Peach developing fruit mesocarp Prunus persica cDNA

ACCESSION BU046755

VERSION BU046755.1 GI:122486832

KEYWORDS EST.

SOURCE prunus persica (peach)

ORGANISM Prunus persica; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.

REFERENCE 1 (bases 1 to 631) Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.

AUTHORS Peach Model Genome for Rosaceae

TITLE Unpublished (2002)

COMMENT Contact: Abbott, A.  
Dept of Genetics and Biochemistry  
Clemson University  
122 Long Hall, Clemson University, Clemson, SC 29634, USA  
Tel: 864 656 3060  
Fax: 864 656 6799  
Email: albert@clemson.edu  
Total High Quality bases = 523

Seq primer: TAATGGACTCATATGATGCG

High quality sequence Stop: 631.

Location/Qualifiers 1..631

FEATURES source

/organism="Prunus persica"

/mol\_type="mRNA"

/db\_xref="taxon:3760"

/clone="PP\_LBA0027104f"

/tissue\_type="Mesocarp"

/lab\_host="B. Coli"

/clone\_lib="Peach developing fruit mesocarp"

/note="Vector: plbluescript II SK(-); Site 1: B'CORI; Site 2: XbaI; authority=Prunus Persica L. Batch; The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis go to http://www.genome.clemson.edu/projects/peach. To order this clone go to http://www.genome.clemson.edu/orders"

ORIGIN

Query Match	8.1%	Score	295.6;	DB	5;	Length	631;
Best Local Similarity	71.7%	Pred.	No.	7.8e-65;			
Matches	407;	Conservative	0;	Mismatches	149;	Indels	12;
						Gaps	1;

Db 1708 ATGATGTTGTTGCTGAGCTTGTGAGCTACCCAAAGCAATTTCAGCAG 1767

Qy 64 ATRACTGCTGCTGAGCTGTTATTGTAGCTACCCCAAGTAATCCCGCAGA 123

Db 1768 TCATTGAGTTAGAAGTGTGACTWGCTCTGTGCTCAAGCTTATTGTAAGAG 1827

Db 124 AGCTTGAGTTATGGGGGGCTCTGGCTTGTGCTAAGCTTAAA 183

Db 1828 CCACACCTTTAAGGAGCTGATAGAACTTCAGCAACTTCAGAGCTAAGTAATG 1887

Db 184 CCTCATCACATTCAAGATCTGAAACCTATTCAGAAGACTCTGACTCTAACA 243

Qy 1888 GCTATGGAGATCTGGCGATGTTGCTGTGTTGATGACAGGAAAGTAGCTCGT 1947

Db 244 GCTGTGAGACATTCCTCTGAGCAACTTATTAACCAAGAAGCAGCTAGACTTGT 303

Qy 1948 CTAGAAAGGGACTCTGTCGACTGCTTAAAGGAAAGTGTGATGATGCCATTGGT 2007

Db 304 TTGGAGAGGGACTCTGTCGACTCTCTCTAGGGACCTGTTGAGCAGCTGTTGGT 363

Qy 2008 GGCTTAGAGCTGAGTGTACATGATGACCATATAGAAATCATCTGTTGAG 2067

Db 364 GGCTTAGAGCTGAGTGTACATGATGACCATATAGAAATCATCTGTTGAG 423

Qy 2068 ATTCAATCTGATG-----ACAATGATGATCTGCAGTCTGTTGAGCTTGGT 2115

Db 424 AACCTCAAGGATGACGATGACAATGACAAATGATGATGATCTCTGGACTTGCAGCA 483

Qy 2116 TTGAAACCTGGTGGCAGGGTGTCTTCCTAGTTGAGCAGACCAAGATAAAA 2175

Db 484 TTGGAGACCTGGTGGAGGTGTTCCCGAGTTGAGAACCAAGAGCTAG 543

Qy 2176 TTAAACTCTGGGACTATGATGATGCTTATGTTGAGTACTCTGAAAGAGTGG 2235

Db 544 TTCAAGCTGGAGACTACATGATGATGATGATCCTACAGCTCTGAGATACTAGAAAGCTGGT 603

Qy 2235 GTAGTCAGGGCTCTCTTAGTGCTG 2263

Db 604 GGCCTAAATGGTCACCCCTTACCTGCTG 631

RESULT 13

LOCUS C0079829 832 bp mRNA linear EST 15-JUN-2004

DEFINITION GR\_Ea2007\_r GR\_Ea Gossypium raimondii cDNA clone GR\_Ea2007\_3', mRNA sequence.

ACCESSION C0079829

VERSION C0079829.1 GI:48749310

KEYWORDS EST.

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 832) Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C., Udall,J.A., Rapp,R., Wendel,J.F., Rao,K., Soderlund,C. and Wing,R.A.

AUTHORS Title Global assembly of Cotton ESTs

COMMENT Unpublished (2004)

Contact: Rod A. Wing  
Arizona Genomics Institute  
The University of Arizona  
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: http://genome.arizona.edu

FEATURES	Plate: 42 row: 0 column: 07.
SOURCE	Location/Qualifiers
Qy	1. /mol_type="mRNA"
	/ab_xref="GR_Ba42007"
	/tissue_type="whole seedlings"
	/dev_stage="first true leaves"
	/lab_host="DH10B"
	/clone_id="GR_Ba"
	/note="Vector: pBMV.SPORT-6.1; Site_1: NotI; Site_2: EcoRV; Library made by Invitrogen with RNA supplied by Wendlie lab. Directional cloned into NotI-EcoRV. Colonies plated/picked by AGI. More glycerol clones held in -80."
ORIGIN	
Query Match	8.0%; Score 293.4; DB 7; Length 832;
Best Local Similarity	69.2%; Pred. No. 3.2e-64;
Matches	424; Conservative 0; Mismatches 171; Indels 18; Gaps 1;
Db	1 trCTGCGAGAAGTTGAGTTATGGAGTTAGCTGAGTTATGGAGTTA 1816
Qy	1817 TGTCTGAGAGCCACACCTTTCAGATGCTATGCAATTCCAGAACTTCAGAG 1876
Db	61 TAACACAGAAACCTCCTCATCAGAGATGCTATAACCTCTTCAGAGCTTCAAGCGA 120
Qy	1877 CTAGGTAATGGGTATGGAGATTCCTGGATGTGTATGATAGAACGGATAATTGGAGA 1936
Db	121 CTAAGGTRACACTCTTGAGAACAGTGTCTCTCTCTGAGCTGAGCTGAGA 180
Qy	1937 TAGACTCGGTCTAGAAGAGGGACTCTGTCAGCTGAGCTATGGCAAATGTGAATGCC 1996
Db	181 TAGACTTGCTTGTGAGAGGGCTCTGTCTATGCTGTGGGGAGCTTGTAGTGCC 240
Qy	1997 GTAGTGCTGTGCTTGTAGACAGGATTGAGATCCAGTATGGCT 2056
Db	241 GTCTGCTGGTTGGCTTAGACAGHTAGCTCCCTTATAGAAATACATCTATGGAT 300
Qy	2057 TTGTTTGGAGAAATTCAATCGTGTAGCATGATCTCCGGACTATGGAAATGT 2116
Db	301 TTGCTCTGAAACTCAAGGTTGAGCTGAGCTGAGATCTCCGGGTTTGCAACTGC 360
Qy	2117 TGGAAACCTGGTGGAGGGTTCTCTGGTTGAGAGCACCAAGATAAAAT 2176
Db	361 TGGAGGATGGCTTAATGGAGGGTTTCTCTGAGTTAGAGAACGACAAAGATAGAT 420
Qy	2177 TTAACCTGGGACTACTATGATGATCTATGCTTGTGAGTTAGCTGAGAGTCAGG 2236
Db	421 TCAAGCTTGGAGATAATTGATGATCTACTCTCTGAGATTTAGAAAGCTTGAGG 480
Qy	2237 TAGTCAGGGTTCTCTAGCTGCTGCTGCAACTATGCAAGGATGGAGGGAG--- 2292
Db	481 GACGAGGGTGTGACCTCTGCTGCAAGGAGCTAGTGGAGCTTCTGAGCTA 540
Qy	2293 -----CATGAAAGCTAGTGTCTGCTGCAAGGAGCTAGTGGAGCTTCTT 2338
Db	541 CTGAGCTCTGTCATGATGAAAGCTAGTGTCTGAGGGCTTCTC 600
Qy	2339 CCCGCTATACAGA 2351
Db	601 TCTGTCGCTCAGA 613
RESULT 14	
CR486014/c	
LOCUS	CR486014 751 bp DNA linear GSS 11-JUN-2004
DEFINITION	Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago truncatula, genomic survey sequence.
ACCESSION	CR486014
VERSION	CR486014.1 GI:48647590
FEATURES	GSS.
SOURCE	Medicago truncatula (barrel medic)
ORGANISM	Medicago truncatula
REFERENCE	1 (bases 1 to 751)
AUTHORS	
TITLE	
JOURNAL	
Submitted (10-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry Cedex - FRANCE (E-mail : seqref@genoscope.cnrs.fr - Web : www.genoscope.cnrs.fr)	
FEATURES	
SOURCE	
ORIGIN	
Query Match	7.2%; Score 265.4; DB 9; Length 751;
Best Local Similarity	64.1%; Pred. No. 5.4e-57;
Matches	476; Conservative 0; Mismatches 216; Indels 51; Gaps 3;
Db	1491 ATGATTCAGCTGGAAAGACTAAATGGTTAGCTGCGGTGCGGGAAATTGGCT 1550
Qy	744 ATGACACATCGAGCGCAGAGAGAGCTCTGCGGGTTGACCGTGTGAGA 685
Db	1551 TTGGAGGGTGGAGCATGCTGCTGTTGGGGTTGACCGTGTGAGAATTGATG 1610
Qy	684 TTGGAGGTGGAGGAGCGAGCAATGGCTGAGGTTACATGATGAGT 625
Db	1611 AGCGGTTTGTGAAATGACAGCTGCTGAGCAGGATAACAGTTAGATA-----CCTT 1664
Qy	624 AGGCATTCCTGCATATGAAAGCTGCGGAACAGGTGTTCTAGAAATATGGCAGTT 565
Db	1665 TTAAATTCTTCTAGATGATAACTTGTGGTTCTCATTTAATGAT----- 1713
Qy	564 TTATTTAATTCTCTGGCCCTTGTGTTGTTGAGCTTACATTATCT 505
Db	1714 -----GTTGCTGTGTTGAGCTTGTGTTGAGCTTACATTATCT 1745
Qy	504 CTTCATTAATCTTTATCTCAAATCTTTTAAAGCTGACATTGTGACAC 445
Db	1746 CCCAASAAATTCCACGAGAGTCATTGAGTTACGAGTTGACTTCTCTCTGGC 1805
Qy	444 ACCTAGGAAATTCACGACTGAGATTTGAGCTTCTGAGCTTCTGAGCT 385
Db	1806 TCAAGCTTTATGGTAAAGAACCCACCTTTACGGAGCTGATAAGGAAATCCGA 1865
Qy	384 ACAGGCTTGTAGTTACGCTTCTGAGCTTCTGAGCTTCTGAGCT 325
Db	1866 ACTTCAGCAGGCTGAGCTATGCTATGGAGATCTCTGCTGTTATGATAC----- 1919
Qy	324 ATTCAACAACTAGTACAACTGAGGATGGCTCCCTCTGTTATCTCCCTGGA 265
Db	1980 CAAGTTGATGATGCGGATCTGGTGGCTTAGACAGTGGAGATCAGATAGGA 2039
Qy	204 GGAGCTTGTGATCATGCTGATCTGGTGGCTGAGCTCAGCTTATGAA 145
Db	264 GATGGAGAGAGAGAGATGTTGAGTTGAGCTGAGCTGAGCTGCTG 205
Qy	2040 TCCAGCTTATGGAGCTTGTGAGATTCAGATGAGATGATGATGATCTCCC 2099
Db	144 CCATCTTATAGACTTATGAGCTTATGGAAAGCCAGAGGTTGAGAGACAGTGTCTC 85

Qy 2100 TGGACTATGCAATTGCTGGAACCTGTTGGAGGGTTGTTCTTAGTCAGAGA 2159  
 Db 84 TGCCTCTGTAAATGTTGAGACATGGTGTGAGGTTCCCTAGTTAGAGA 25  
 Qy 2160 CACCAAGATAAAATTAAAC 2182  
 Db 24 TACTAAGGACAACTTAAAC 2  
 RESULT 15  
 CO11747 LOCUS CO11747 GR\_Eb019K01\_r GR\_Eb Gossypium raimondii EST 16-JUN-2004  
 DEFINITION 849 bp mRNA linear EST 16-JUN-2004  
 ACCESSION 3' mRNA Sequence.  
 VERSION CO11747.1 GI:48815734  
 KEYWORDS EST.  
 ORGANISM Gossypium raimondii  
 Spermatophyta; Embryophyta; Tracheophyta;  
 Rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
 REFERENCE 1 (bases 1 to 849)  
 AUTHORS Kim, Y.; Kudrna, D.; Hatfield, J.; Stum, D.; Mueller, C.;  
 Udali, J.A.; Rapp, R.A.; Wendel, J.F.; Rao, K.; Soderlund, C. and  
 Wing, R.A.  
 TITLE Global assembly of Cotton ESTs  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Rod A. Wing  
 Arizona Genomics Institute  
 The University of Arizona  
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
 Tel: 520 626 9595  
 Fax: 520 621 1259  
 Email: <http://genome.arizona.edu>  
 Plate: 019 row: K column: 01.  
 FEATURES Location/Qualifiers  
 Source 1. .849 /organism="Gossypium raimondii"  
 /mol\_type="mRNA"  
 /db\_xref="Raxom:29730"  
 /clone="GR\_Eb019K01"  
 /tissue\_type="floral"  
 /dev\_stage="3 to +3 DPA"  
 /lab\_host="DHL10B"  
 /clone\_lib="GR\_Eb"  
 /note="Vector: pCMV.SPORT-6.1; Site1: NotI; Site2:  
 ECORV; Library made by Invitrogen with RNA supplied by  
 Wendle lab. Directional cloned into NotI-EV. Colonies  
 plated/picked by AGI. More glycerol clones held in -80."  
 ORIGIN  
 Query Match 6.9%; Score 254; DB 7; Length 849;  
 Best Local Similarity 58.7%; Prod. No. 5e-54; Mismatches 325; Indels 39; Gaps 3;  
 Matches 518; Conservative 0; Mismatches 325; Indels 39; Gaps 3;  
 Qy 1828 CCACACCTTTAGGAGCTGTAAGGAACTTACGCAACTTACGAGGCTTAAGGTTAG 1887  
 Db 1 CCTCATCTCATAGAGATCTGTATAACCTCTTCAGCAGCTTGAGAATAGTAACA 60  
 Qy 1888 GCTATGGAGATCTGGATGTTATGACCGATAATGGAGATAGCTCGT 1947  
 Db 61 ACTCTTGAGAACAGTGCTCTCTATGCTCGTACCGTAACTGAGATAGCTCT 120  
 Qy 1948 CTAGAAAGGGGACTCTGCACTCTATAGGAAAGTGTGATGGTATGGGTG 2007  
 Db 121 TTGGAGAGGGCTCTGTCATCTCTGGGGAGCTGTGATGAGTGGTGGTGGT 180  
 Qy 2008 GGCTTAGAGCTGGATTCACTATAGGAACTCAGTTATGGAGCTGGTGTGTTGGAG 2067  
 Db 181 GGCTTAGAGCTGGATTCACTATAGGAACTCAGTTATGGAGCTGGTGTGTTGGAG 240  
 Qy 2068 ATTCAAATGCTGATGACATGATGATCTCCCTGACTATGCAATTGTTGGAAACCTGG 2127

Db 241 AACTCAGGATGAGATCACAGAGATCTCCCGGCTTGCAACTCTGGGGATGG 300  
 Db 2128 TTGGCAGGGTCTCTCTCTAGTTCAGAGAACAGATAAAATTAAACTCGG 2187  
 Qy 301 CTATGGAGGTGAGTTTCCTGATGAGATTTGAGACAGATACATTCAGCTTGA 360  
 Db 2188 GACTACTATGATGCCATGGTTGAGTACTTGGAGAGTGGAGGTTCAGGGT 2247  
 Qy 361 GATTTATGATGATCCCTACTGTCCTGAGATATTGAGAAGTGGAGGAGGGT 420  
 Db 2248 TCTCTTGTAGTGTGTCACACTATGGCAAGGATTGGCCAGCATGTGAAGCTAGT 2307  
 Db 421 TCAACCTCTGCTGAGCCCGAGCTATAGTGAGGATAGSTGAGGCTACTGGAGTCT 480  
 Qy 2308 GCTATGAGGCACTGAGAAAGTTCTTCCCGTATACAGATAGAACTGGGTGA 2367  
 Db 481 GAT-----CAGTAAAGGGCTATGTCATTG 507  
 Qy 2368 CCCAGGATGTCAGAGACAGTTTGTGAGATCCCTGTTGTAACATGTGGCCG 2427  
 Db 508 GCATGAGGTTCTCTCTGTCAGAGACATCAATTAGAT 567  
 Qy 2428 GATGTGAGCTGGCTTTATGCAAGAGCTGTAACCCCTGTGAAACTTGAAC 2487  
 Db 568 GGTGAGATGAACTTCTCTGTGAAAGTGGAGGACTCTGGAAACCTGATCA 627  
 Qy 2438 AATGATTAGCAATTGAGCTGGCTGGGTCTC--AGAGAGTAGCTGTGAAACTACTGT 2544  
 Db 628 GAGCATCTGCCATCTACTGAGTGTCTGGATATAGGAGCTTGGAAATGATCA 687  
 Qy 2545 GAATGTGCGTGTGATGTTAAGGGGCACTGTAAGATCTTACGTGGTGGT 2604  
 Db 688 GAGAAACATTAAGTGAACATTAAAGATGCGAGTCAGATCATCTCTGTTG 747  
 Qy 2605 GCAATTGGACTGTTCACTCTGTCAGCAGAACTATT---TCTAAAGCACCTCACT 2661  
 Db 748 GIAATGGGTGATGACITGGTGGCTGAGTTTACTGTAATTAGTCATCA 807  
 Qy 2662 TTTCACCCAGGATATGTTCTCTATGGATCTGATGTC 2703  
 Db 808 GTGACCGGTAAGGCATCGTCCAGCTTGGCACTGATGTC 849  
 Search completed: June 10, 2005, 00:53:15  
 Job time : 11203 secs

THIS PAGE BLANK (USPTO)